Unit C3

Linear transformations

Introduction

In this unit you will study functions between vector spaces. You will begin by looking more closely at some particular functions that have $\mathbb{R}^2$ as their domain and codomain, such as rotations and reflections. These functions map parallel lines to parallel lines, preserve scalar multiples and map the zero vector to itself. Algebraically, these functions preserve the operations of addition and scalar multiplication in the vector space $\mathbb{R}^2$. There is a special name for functions that preserve addition and scalar multiplication between vector spaces: they are called *linear transformations.* You will see that such functions have a matrix representation. This link between linear transformations and matrices enables us to relate the properties of matrices with those of linear transformations. Finally, you will meet an important result concerning linear transformations, known as the *Dimension Theorem.* This theorem has a number of consequences. For example, it enables us to show how the number of solutions of a system of m linear equations in n unknowns depends on the values of m and n.

Many results from Units C1 *Linear equations and matrices* and C2 *Vector spaces* are used in this unit; so make sure that you understand the main ideas of those units before starting your study of this one.

1 Introducing linear transformations

In this section you will see that we can generalise properties of functions that have $\mathbb{R}^2$ as their domain and codomain to functions between other vector spaces.

1.1 What is a linear transformation?

We begin by investigating the properties of some simple but important functions, often called transformations, which map the vector space $\mathbb{R}^2$ to itself. For each one, a diagram shows the effect of the transformation on the square whose corners are at $(0,0)$, $(0,1)$, $(1,1)$ and $(1,0)$, and the effect on the vector $(1,1)$; part of the square is shaded for clarity.

We will investigate the following four functions: dilation, scaling, rotation and reflection.

For any real number k, a k-**dilation** of $\mathbb{R}^2$ scales (or stretches) vectors by a factor k with respect to the origin.

When $k = 2$, the magnitude of a vector is doubled, as illustrated in Figure 1.

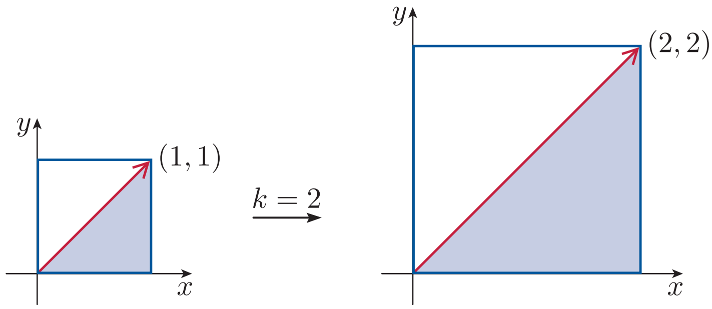


Figure 1 A 2-dilation

When $k = \frac{1}{2}$, the magnitude of a vector is halved, as illustrated in Figure 2.

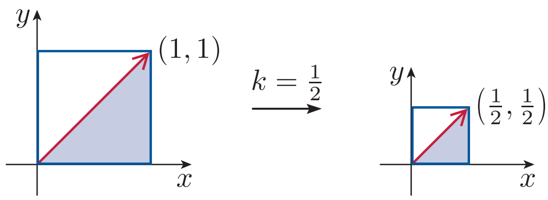


Figure 2 A $\frac{1}{2}$-dilation

When k is negative, the direction of a vector is reversed – as illustrated in Figure 3 for the case $k = -2$.

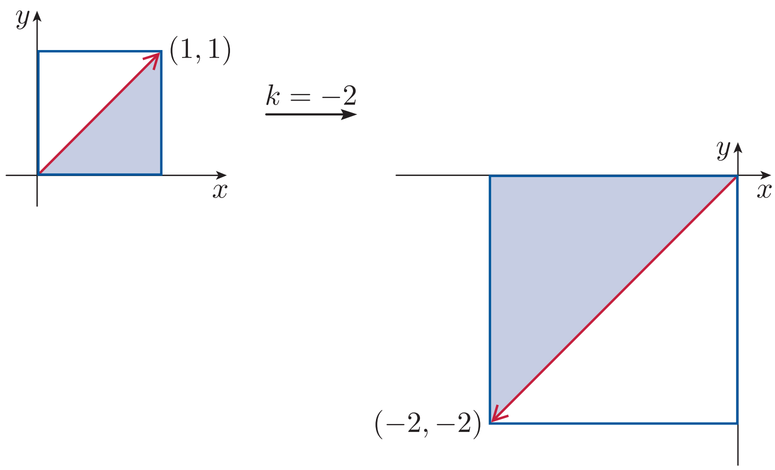


Figure 3 A -2-dilation

For any real numbers k and l, a (k, l)-**scaling** of $\mathbb{R}^2$ scales vectors by a factor k in the x-direction and by a factor l in the y-direction. Figure 4 shows the effect of a $\left(2, \frac{1}{2}\right)$-scaling.

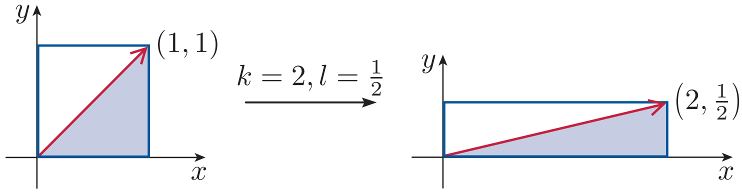


Figure 4 A $\left(2, \frac{1}{2}\right)$-scaling

Figure 5 shows the effect of a $(-1, 3)$-scaling.

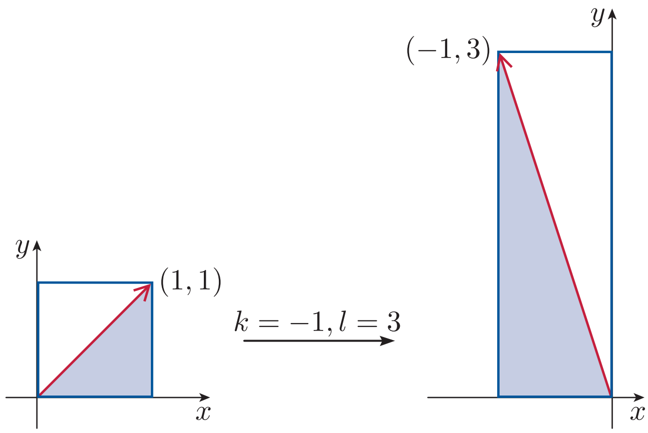


Figure 5 A $(-1, 3)$-scaling

A **rotation** r_θ of $\mathbb{R}^2$ rotates vectors anticlockwise through an angle θ about the origin $(0, 0)$.

Figure 6 shows the effect of a rotation $r_{\pi/4}$.

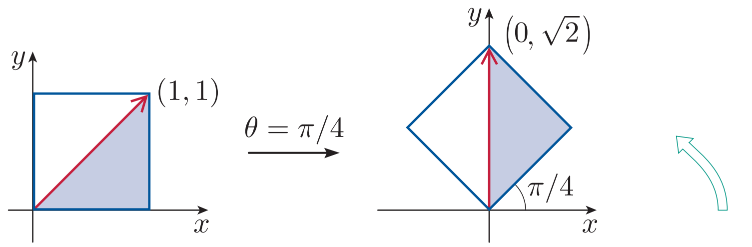


Figure 6 A rotation $r_{\pi/4}$

Figure 7 shows the effect of a rotation $r_{\pi/2}$.

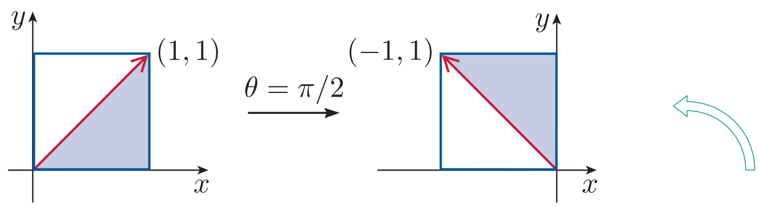


Figure 7 A rotation $r_{\pi/2}$

A **reflection** q_ϕ of $\mathbb{R}^2$ reflects vectors in the straight line through the origin that makes an angle ϕ with the x-axis (measured anticlockwise).

Figure 8 shows the effect of reflection $q_{\pi/4}$.

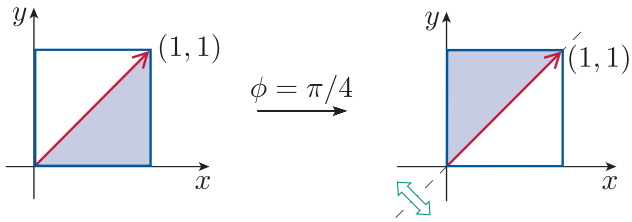


Figure 8 A reflection $q_{\pi/4}$

Figure 9 shows the effect of a reflection $q_{\pi/2}$.

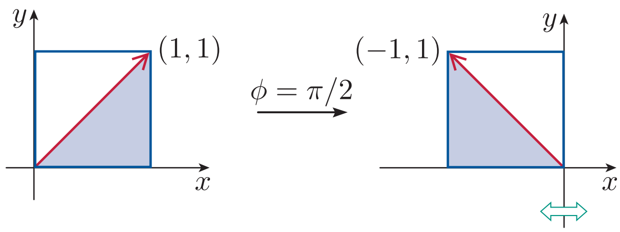


Figure 9 A reflection $q_{\pi/2}$

Exercise C82

For each of the following functions, draw a diagram to show the effect of the function on the rectangle with corners at $(0,0)$, $(2,0)$, $(2,1)$ and $(0,1)$, and on the vector $(2,1)$. State whether the function is a dilation, a scaling, a rotation or a reflection.

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x,y) \longmapsto (2x, 3y)$

(b) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x,y) \longmapsto (x, -y)$

(c) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x,y) \longmapsto (-y, x)$

We now use matrix multiplication, from Unit C1, to obtain algebraic definitions of the four types of function defined geometrically above: dilation, scaling, rotation and reflection.

A k-dilation of $\mathbb{R}^2$ maps (x,y) to (kx,ky). This can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix}.$$

A (k,l)-scaling of $\mathbb{R}^2$ maps (x,y) to (kx,ly). This can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ly \end{pmatrix}.$$

An algebraic definition for a rotation r_θ of $\mathbb{R}^2$ can be obtained by considering Figure 10, where r_θ maps (x, y) to (x', y').

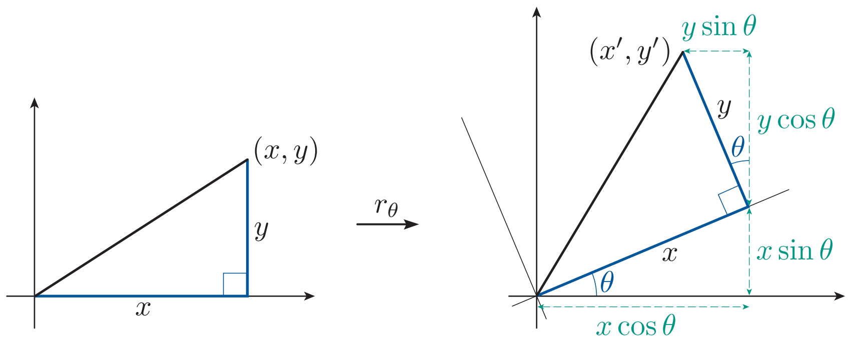


Figure 10 A rotation r_θ (through an angle of θ).

It can be seen that

$$(x, y) \longmapsto (x', y') = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

This can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

For example, $r_{\pi/6}$ can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2}x - \frac{1}{2}y \\ \frac{1}{2}x + \frac{\sqrt{3}}{2}y \end{pmatrix}.$$

Similarly, it can be shown (you will show this in Exercise C88) that a reflection q_ϕ of $\mathbb{R}^2$ can be defined algebraically by

$$(x, y) \longmapsto (x\cos 2\phi + y\sin 2\phi, x\sin 2\phi - y\cos 2\phi).$$

This can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos 2\phi + y\sin 2\phi \\ x\sin 2\phi - y\cos 2\phi \end{pmatrix}.$$

For example, $q_{\pi/6}$ can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x + \frac{\sqrt{3}}{2}y \\ \frac{\sqrt{3}}{2}x - \frac{1}{2}y \end{pmatrix}.$$

We have seen that each of the four types of function can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

for some real numbers a, b, c and d.

The existence of a matrix representation is not the only property shared by these functions of the plane: they also share several striking geometric properties. In each of the examples, the image of the unit square is either a square or a rectangle; each of these functions maps straight lines to straight lines – indeed, each maps parallel lines to parallel lines. Any function that maps parallel lines to parallel lines will map parallelograms to parallelograms. Another geometric property shared by these four functions is that they also all map the origin to itself.

Figure 11 shows the effect of a general transformation t on two vectors, $\mathbf{v}_1$ and $\mathbf{v}_2$, where t maps parallelograms to parallelograms and preserves the origin.

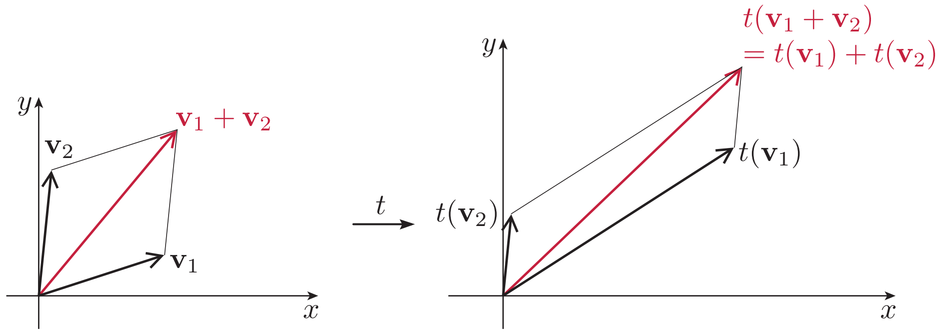


Figure 11 Parallelograms are mapped to parallelograms

Bearing in mind the Parallelogram Law for addition of vectors from Unit A1 *Sets, functions and vectors*, this illustrates that for each function t in one of the four classes above, we have

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2.$$

Such a function t also preserves scalar multiples, as illustrated in Figure 12; that is, if $\alpha\mathbf{v}$ is a scalar multiple of a vector $\mathbf{v}$, then the image of $\alpha\mathbf{v}$ under t is a scalar multiple of the image of $\mathbf{v}$ under t.

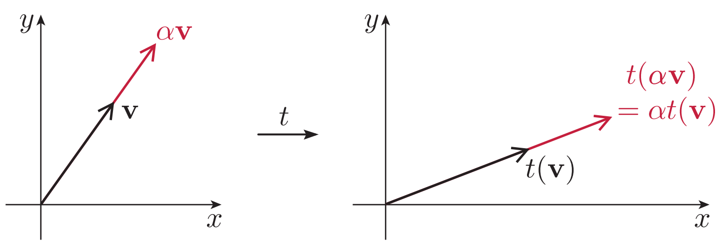


Figure 12 Scalar multiples are preserved

We have

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in \mathbb{R}^2,\ \alpha \in \mathbb{R}.$$

We use these two algebraic properties to define a *linear transformation* from any vector space to another: a linear transformation is any function from a vector space V to a vector space W that has these two algebraic properties. You will see why these functions are called *linear* transformations in Subsection 1.3.

Definition

Let V and W be vector spaces. A function $t : V \longrightarrow W$ is a **linear transformation** if it satisfies the following properties.

LT1 $t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2)$, for all $\mathbf{v}_1, \mathbf{v}_2 \in V$.

LT2 $t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v})$, for all $\mathbf{v} \in V$, $\alpha \in \mathbb{R}$.

In Section 2 we show that the functions between finite-dimensional vector spaces that have these two properties are precisely those functions that have matrix representations.

Suppose that $t : V \longrightarrow W$ is a linear transformation. It follows from property LT1 that if we know the images of two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ under t, then we can find the image of the vector $\mathbf{v}_1 + \mathbf{v}_2$. It follows from property LT2 that if we know the image of a vector $\mathbf{v}$ under t, then we can find the image of any scalar multiple of $\mathbf{v}$.

Thus, once we know the images of some vectors, we can find the images of more vectors by applying properties LT1 and LT2. In fact, if we know the image of each vector in a *basis* for V, then we can find the image of *every* vector in V. It is this property that makes linear transformations so important; we will prove it at the end of this section.

All the functions of the plane that we have studied map the origin to itself. In fact, any linear transformation $t : V \longrightarrow W$ maps the zero vector of V to the zero vector of W. To see this, we use property LT2:

$$t(\mathbf{0}) = t(0\mathbf{0}) = 0t(\mathbf{0}) = \mathbf{0}.$$

We have proved the following result.

Theorem C37

Let $t : V \longrightarrow W$ be a linear transformation. Then $t(\mathbf{0}) = \mathbf{0}$.

It follows from Theorem C37 that a function $t : V \longrightarrow W$ where $t(\mathbf{0}) \neq \mathbf{0}$ is *not* a linear transformation; for example, the function

$$\begin{aligned} t &: \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (y - 1, x) \end{aligned}$$

is not a linear transformation because

$$t(\mathbf{0}) = t(0, 0) = (-1, 0) \neq \mathbf{0}.$$

However, a function t with the property $t(\mathbf{0}) = \mathbf{0}$ is not necessarily a linear transformation. For example, the function

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, |y|) \end{aligned}$$

satisfies $t(\mathbf{0}) = \mathbf{0}$ but is not a linear transformation. To see this, consider the two vectors $(0, 1)$ and $(0, -1)$. LT1 is not satisfied because

$$t\left((0, 1) + (0, -1)\right) = t(0, 0) = (0, 0)$$

and

$$t(0, 1) + t(0, -1) = (0, 1) + (0, 1) = (0, 2).$$

LT2 is also not satisfied; this can be shown, for example, by taking the vector $(0, 1)$ and $\alpha = -1$.

The following strategy can be used to test whether a given function is a linear transformation.

Strategy C14

To determine whether or not a given function $t : V \longrightarrow W$ is a linear transformation, do the following.

1. Check whether $t(\mathbf{0}) = \mathbf{0}$; if not, then t is not a linear transformation.
2. Check whether t satisfies the following two properties.

   LT1 $\quad t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad$ for all $\mathbf{v}_1, \mathbf{v}_2 \in V$.

   LT2 $\quad t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad$ for all $\mathbf{v} \in V$, $\alpha \in \mathbb{R}$.

The function t is a linear transformation if and only if both these properties are satisfied.

You may have noticed that if the two properties in step 2 of the strategy both hold, then t is a linear transformation and we do not also need to check step 1. We have, however, included step 1 in the strategy as this can provide a quick way of showing that some functions are *not* linear transformations. On the other hand, if step 1 holds but either one of properties LT1 or LT2 fails, then you do not need to check the other.

Worked Exercise C49

Use Strategy C14 to determine whether or not each of the following functions is a linear transformation.

(a) $\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (2x, y) \end{aligned}$

(b) $\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto ((x + y)^2, y^2) \end{aligned}$

Solution

(a) You may notice that t is a $(2,1)$-scaling, and so expect it to be a linear transformation.

Here $t(\mathbf{0}) = \mathbf{0}$, so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2.$$

In $\mathbb{R}^2$, let $\mathbf{v}_1 = (x_1, y_1)$ and $\mathbf{v}_2 = (x_2, y_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2) \\ &= (2(x_1 + x_2), y_1 + y_2) \\ &= (2x_1 + 2x_2, y_1 + y_2) \end{aligned}$$

and

$$\begin{aligned} t(\mathbf{v}_1) + t(\mathbf{v}_2) &= (2x_1, y_1) + (2x_2, y_2) \\ &= (2x_1 + 2x_2, y_1 + y_2). \end{aligned}$$

These expressions are equal, so LT1 is satisfied.

Finally, we check whether t satisfies LT2:

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in \mathbb{R}^2,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} = (x, y)$ be a vector in $\mathbb{R}^2$ and let $\alpha \in \mathbb{R}$. Then

$$t(\alpha\mathbf{v}) = t(\alpha x, \alpha y) = (2\alpha x, \alpha y)$$

and

$$\alpha\, t(\mathbf{v}) = \alpha\, t(x, y) = \alpha(2x, y) = (2\alpha x, \alpha y).$$

These expressions are equal, so LT2 is satisfied.

Since LT1 and LT2 are satisfied, t is a linear transformation.

(b) Here $t(\mathbf{0}) = \mathbf{0}$, so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2.$$

In $\mathbb{R}^2$, let $\mathbf{v}_1 = (x_1, y_1)$ and $\mathbf{v}_2 = (x_2, y_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2) \\ &= ((x_1 + x_2 + y_1 + y_2)^2, (y_1 + y_2)^2) \end{aligned}$$

and

$$\begin{aligned} t(\mathbf{v}_1) + t(\mathbf{v}_2) &= ((x_1 + y_1)^2, y_1^2) + ((x_2 + y_2)^2, y_2^2) \\ &= ((x_1 + y_1)^2 + (x_2 + y_2)^2, y_1^2 + y_2^2). \end{aligned}$$

These expressions are not equal in general, so LT1 is not satisfied.

Thus t is not a linear transformation.

Since property LT1 is not satisfied, there is no need to check property LT2; however, in this case it also does not hold.

Exercise C83

Use Strategy C14 to determine whether or not each of the following functions is a linear transformation.

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x + 3y, y)$

(b) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x + 2, y + 1)$

In Exercise C83(a) you showed that the function

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (x + 3y, y)$$

is a linear transformation. This function is an example of a *shear*, or *skew*, of $\mathbb{R}^2$.

As illustrated in Figure 13, in general, a **shear** of $\mathbb{R}^2$ in the x-direction by a factor k is the linear transformation

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (x + ky, y).$$

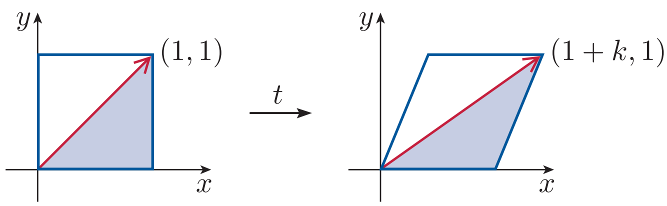


Figure 13 A shear in the x-direction by a factor of k

In Exercise C83(b) you showed that the function

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (x + 2, y + 1)$$

is not a linear transformation. This function is an example of a *translation* of $\mathbb{R}^2$.

As illustrated in Figure 14, in general, a **translation** of $\mathbb{R}^2$ by (a, b) is the function

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (x + a, y + b).$$

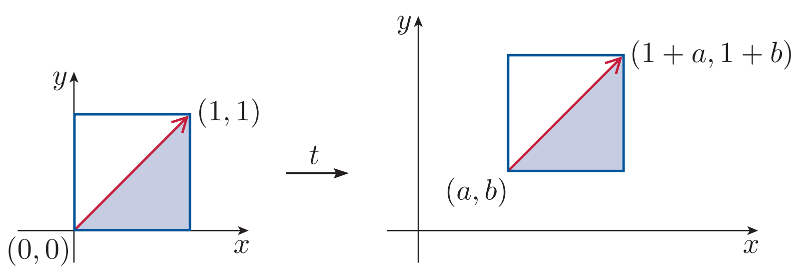


Figure 14 A translation by (a, b)

A translation is not a linear transformation unless $a = b = 0$, since otherwise it does not map the origin to itself.

1.2 Examples of linear transformations

You have seen many examples of functions from $\mathbb{R}^2$ to $\mathbb{R}^2$. In general, given any two vector spaces V and W, we can define functions from V to W. For example, consider the function t from $\mathbb{R}^3$ to $\mathbb{R}^2$ that projects each vector in $\mathbb{R}^3$ onto the (x, y)-plane, as illustrated in Figure 15:

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x, y). \end{aligned}$$

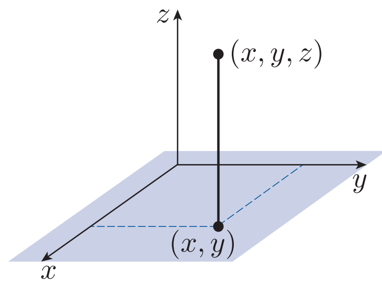


Figure 15 A projection from $\mathbb{R}^3$ onto the (x, y)-plane

This function is a linear transformation as shown in the next worked exercise.

Worked Exercise C50

Show that the following function t from $\mathbb{R}^3$ to $\mathbb{R}^2$ is a linear transformation.

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x, y) \end{aligned}$$

Solution

Note that the question says 'show', not 'determine'; we know that it *is* a linear transformation. Thus we use the definition rather than Strategy C14 and avoid the need to check whether $t(\mathbf{0}) = \mathbf{0}$.

First we show that t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3.$$

In $\mathbb{R}^3$, let $\mathbf{v}_1 = (x_1, y_1, z_1)$ and $\mathbf{v}_2 = (x_2, y_2, z_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_1 + x_2, y_1 + y_2) \end{aligned}$$

and

$$\begin{aligned} t(\mathbf{v}_1) + t(\mathbf{v}_2) &= t(x_1, y_1, z_1) + t(x_2, y_2, z_2) \\ &= (x_1, y_1) + (x_2, y_2) \\ &= (x_1 + x_2, y_1 + y_2). \end{aligned}$$

These expressions are equal, so LT1 is satisfied.

Next we show that t satisfies LT2:

$$t(\alpha \mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in \mathbb{R}^3,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} = (x, y, z)$ be a vector in $\mathbb{R}^3$ and let $\alpha \in \mathbb{R}$. Then

$$t(\alpha\mathbf{v}) = t(\alpha x, \alpha y, \alpha z) = (\alpha x, \alpha y)$$

and

$$\alpha\, t(\mathbf{v}) = \alpha\, t(x, y, z) = \alpha(x, y) = (\alpha x, \alpha y).$$

These expressions are equal, so LT2 is satisfied.

Since LT1 and LT2 are satisfied, t is a linear transformation.

Worked Exercise C51

Determine whether or not the following function is a linear transformation.

$$\begin{aligned} t : \mathbb{R}^4 &\longrightarrow \mathbb{R}^2 \\ (x, y, z, w) &\longmapsto (xy, z) \end{aligned}$$

Solution

The question says 'determine', so here we do use the strategy.

We use Strategy C14.

Since $t(\mathbf{0}) = \mathbf{0}$, t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^4.$$

In $\mathbb{R}^4$, let $\mathbf{v}_1 = (x_1, y_1, z_1, w_1)$ and $\mathbf{v}_2 = (x_2, y_2, z_2, w_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2) \\ &= ((x_1 + x_2)(y_1 + y_2), z_1 + z_2) \end{aligned}$$

and

$$\begin{aligned} t(\mathbf{v}_1) + t(\mathbf{v}_2) &= (x_1y_1, z_1) + (x_2y_2, z_2) \\ &= (x_1y_1 + x_2y_2, z_1 + z_2). \end{aligned}$$

Since $(x_1 + x_2)(y_1 + y_2) \neq x_1y_1 + x_2y_2$ in general, LT1 is not satisfied.

Thus t is not a linear transformation.

Exercise C84

Determine whether or not each of the following functions is a linear transformation.

(a) $\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^4 \\ (x, y) &\longmapsto (x, y, x, y) \end{aligned}$

(b) $\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto x^2 \end{aligned}$

(c) $\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^4 \\ (x, y, z) &\longmapsto (x, y, z, 1) \end{aligned}$

In the previous subsection we gave an algebraic definition of a rotation of $\mathbb{R}^2$. Similarly, a rotation of $\mathbb{R}^3$ in an anticlockwise direction about the z-axis through an angle θ, as illustrated in Figure 16, is given by

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z). \end{aligned}$$

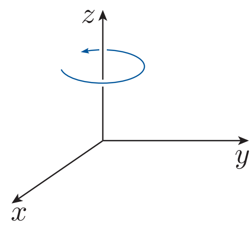


Figure 16 A rotation about the z-axis through an angle θ

Exercise C85

Show that the following function t is a linear transformation.

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z). \end{aligned}$$

So far we have considered functions $t : V \longrightarrow W$ where $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ for some $m, n \in \mathbb{N}$. There are, however, many functions between other types of vector space.

Recall from Unit C2 that the vector space P_n is the set of all polynomials of degree less than n, so

$$\begin{aligned} P_3 &= \{p(x) : p(x) = a + bx + cx^2,\ a, b, c \in \mathbb{R}\}, \\ P_2 &= \{p(x) : p(x) = a + bx,\ a, b \in \mathbb{R}\}. \end{aligned}$$

Worked Exercise C52

Consider the function that maps each polynomial $p(x) = a + bx + cx^2$ in P_3 to its derivative $p'(x) = b + 2cx$ in P_2:

$$\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x). \end{aligned}$$

Determine whether or not this function is a linear transformation.

Solution

We use Strategy C14.

Since the zero element of P_3 is $p(x) = 0$, we have $p'(x) = 0$ and thus $t(\mathbf{0}) = \mathbf{0}$; so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(p(x) + q(x)) = t(p(x)) + t(q(x)), \quad \text{for all } p(x), q(x) \in P_3.$$

Let $p(x), q(x) \in P_3$. Then

$$t(p(x) + q(x)) = (p(x) + q(x))' = p'(x) + q'(x)$$

and

$$t(p(x)) + t(q(x)) = p'(x) + q'(x).$$

These expressions are equal, so LT1 is satisfied.

Finally, we check whether t satisfies LT2:

$$t(\alpha p(x)) = \alpha\, t(p(x)), \quad \text{for all } p(x) \in P_3,\ \alpha \in \mathbb{R}.$$

Let $p(x) \in P_3$ and $\alpha \in \mathbb{R}$. Then

$$t(\alpha p(x)) = (\alpha p(x))' = \alpha p'(x)$$

and

$$\alpha\, t(p(x)) = \alpha p'(x).$$

These expressions are equal, so LT2 is satisfied.

Since LT1 and LT2 are satisfied, t is a linear transformation.

Exercise C86

Consider the function t from P_3 to itself obtained by adding to each polynomial $p(x) = a + bx + cx^2$ in P_3 the number $p(2) = a + 2b + 4c$:

$$\begin{aligned} t : P_3 &\longrightarrow P_3 \\ p(x) &\longmapsto p(x) + p(2). \end{aligned}$$

Determine whether or not this function is a linear transformation.

There are also linear transformations of infinite-dimensional vector spaces. For example, let V be the vector space of all real functions. An argument similar to that in the solution to Exercise C86 shows that the following function is a linear transformation:

$$\begin{aligned} t : V &\longrightarrow V \\ f(x) &\longmapsto f(x) + f(2). \end{aligned}$$

Zero transformation

Since every vector space contains a zero vector, given any two vector spaces V and W, there is a particularly simple function mapping each vector in V to the zero vector in W:

$$\begin{aligned} t : V &\longrightarrow W \\ \mathbf{v} &\longmapsto \mathbf{0}. \end{aligned}$$

This function is a linear transformation. To show this, we first show that t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in V.$$

Let $\mathbf{v}_1, \mathbf{v}_2 \in V$. Then $\mathbf{v}_1 + \mathbf{v}_2$ is also in V, so

$$t(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{0}$$

and

$$t(\mathbf{v}_1) + t(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

So LT1 is satisfied.

Next we show that t satisfies LT2:

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in V,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} \in V$ and $\alpha \in \mathbb{R}$. Then $\alpha\mathbf{v}$ is also in V, so

$$t(\alpha\mathbf{v}) = \mathbf{0}$$

and

$$\alpha\, t(\mathbf{v}) = \alpha\mathbf{0} = \mathbf{0}.$$

So LT2 is satisfied.

Since LT1 and LT2 are satisfied, t is a linear transformation.

Definition

The **zero transformation** from V to W is the linear transformation

$$\begin{aligned} t : V &\longrightarrow W \\ \mathbf{v} &\longmapsto \mathbf{0}. \end{aligned}$$

Identity transformation

Given a vector space V, there is another particularly simple function, this time from V to itself, mapping each vector in V to itself:

$$\begin{aligned} i_V : V &\longrightarrow V \\ \mathbf{v} &\longmapsto \mathbf{v}. \end{aligned}$$

Exercise C87

Show that the function i_V is a linear transformation.

Definition

The **identity transformation** of V is the linear transformation

$$\begin{aligned} i_V : V &\longrightarrow V \\ \mathbf{v} &\longmapsto \mathbf{v}. \end{aligned}$$

We omit the subscript V when the vector space is clear from the context.

1.3 Linear combinations of vectors

Recall from Unit C2 that a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is an expression of the form $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$, where $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. We end this section by proving that linear combinations of vectors are preserved under a linear transformation; that is, if $\mathbf{v}$ is a given linear combination of vectors $\mathbf{v}_i$, then the image of $\mathbf{v}$ is the same linear combination of the images of the vectors $\mathbf{v}_i$. This explains why these functions are called linear transformations. In fact, some texts use this theorem as the definition of a linear transformation.

Theorem C38

A function $t : V \longrightarrow W$ is a linear transformation if and only if it satisfies

LT3 $t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2)$,

for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and all $\alpha_1, \alpha_2 \in \mathbb{R}$.

Proof We start by proving the 'only if' part using LT1 and LT2 to show that a linear transformation satisfies LT3.

If a function $t : V \longrightarrow W$ is a linear transformation, then it satisfies LT1 and LT2. We show that this implies that it satisfies LT3.

Let $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $\alpha_1, \alpha_2 \in \mathbb{R}$. Then it follows from LT1 that

$$t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = t(\alpha_1\mathbf{v}_1) + t(\alpha_2\mathbf{v}_2),$$

and from LT2 that

$$t(\alpha_1\mathbf{v}_1) + t(\alpha_2\mathbf{v}_2) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2).$$

So t satisfies the property LT3:

$$t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2),$$
$$\text{for all } \mathbf{v}_1, \mathbf{v}_2 \in V \text{ and all } \alpha_1, \alpha_2 \in \mathbb{R}.$$

We now prove the 'if' part using property LT3 to show that LT1 and LT2 are satisfied.

Suppose that a function $t : V \longrightarrow W$ satisfies property LT3. Then it also satisfies LT1 and LT2, since

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in V,$$

is a special case of LT3 with $\alpha_1 = \alpha_2 = 1$, and

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in V,\ \alpha \in \mathbb{R},$$

is a special case of LT3 with $\mathbf{v}_2 = \mathbf{0}$, $\mathbf{v}_1 = \mathbf{v}$, $\alpha_1 = \alpha$ and $\alpha_2 = 0$.

Thus a function is a linear transformation if and only if it satisfies property LT3. ■

We now prove that linear combinations of any number of vectors are preserved under a linear transformation.

Theorem C39

Let $t : V \longrightarrow W$ be a linear transformation. Then

$$t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2) + \cdots + \alpha_n t(\mathbf{v}_n),$$

for all $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ and all $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, $n \in \mathbb{N}$.

Proof We use proof by mathematical induction as in Unit A3 *Mathematical language and proof* and start by writing out clearly what we take $P(n)$ to be.

Let $P(n)$ be the statement

$$t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2) + \cdots + \alpha_n t(\mathbf{v}_n),$$
$$\text{for all } \mathbf{v}_1, \ldots, \mathbf{v}_n \in V \text{ and all } \alpha_1, \ldots, \alpha_n \in \mathbb{R}.$$

Next, we carry out step 1; that is, we check that $P(1)$ holds.

Since t is a linear transformation, LT2 is satisfied, so

$$t(\alpha_1\mathbf{v}_1) = \alpha_1 t(\mathbf{v}_1), \quad \text{for all } \mathbf{v}_1 \in V,\ \alpha_1 \in \mathbb{R}.$$

Thus $P(1)$ is true.

Now we proceed with step 2. We start by stating clearly our assumption, $P(k)$.

We assume that $P(k)$ is true for some positive integer k; that is,

$$t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2) + \cdots + \alpha_k t(\mathbf{v}_k),$$
$$\text{for all } \mathbf{v}_1, \ldots, \mathbf{v}_k \in V \text{ and all } \alpha_1, \ldots, \alpha_k \in \mathbb{R}.$$

We state clearly our desired conclusion, $P(k+1)$.

We wish to deduce that $P(k+1)$ is true; that is,

$$\begin{aligned} &t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k + \alpha_{k+1}\mathbf{v}_{k+1}) \\ &\quad = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2) + \cdots + \alpha_k t(\mathbf{v}_k) + \alpha_{k+1} t(\mathbf{v}_{k+1}). \end{aligned}$$

Now, $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1} \in V$ and all $\alpha_1, \ldots, \alpha_{k+1} \in \mathbb{R}$. We have

$$\begin{aligned} &t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k + \alpha_{k+1}\mathbf{v}_{k+1}) \\ &\quad = t((\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k) + \alpha_{k+1}\mathbf{v}_{k+1}) \\ &\quad = t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k) + t(\alpha_{k+1}\mathbf{v}_{k+1}) \quad \text{(by LT1)} \\ &\quad = t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k) + \alpha_{k+1} t(\mathbf{v}_{k+1}) \quad \text{(by LT2)} \\ &\quad = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2) + \cdots + \alpha_k t(\mathbf{v}_k) + \alpha_{k+1} t(\mathbf{v}_{k+1}) \quad \text{(by } P(k)\text{)}. \end{aligned}$$

We have proved that $P(k) \Rightarrow P(k+1)$.

Thus $P(k) \Rightarrow P(k+1)$, for $k = 1, 2, \ldots$.

Hence, by the Principle of Mathematical Induction, $P(n)$ is true for all $n \in \mathbb{N}$. ■

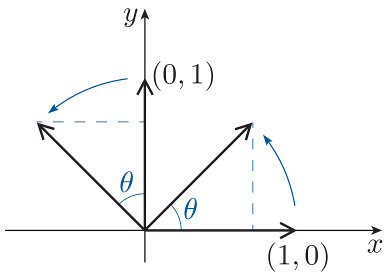


Figure 17 The rotation r_θ

Theorem C39 is an important result. It means that, given a linear transformation $t : V \longrightarrow W$ and the images of each of the vectors in a basis for V, we can determine the image of *any* vector in V.

Consider the linear transformation r_θ that rotates each vector in $\mathbb{R}^2$ anticlockwise through an angle θ about the origin, as illustrated in Figure 17. The standard basis for $\mathbb{R}^2$ is $\{(1,0),(0,1)\}$. From Figure 17, we can check that

$$r_\theta(1,0) = (\cos\theta, \sin\theta) \quad \text{and} \quad r_\theta(0,1) = (-\sin\theta, \cos\theta).$$

We now write each vector (x,y) in $\mathbb{R}^2$ in the form

$$(x,y) = x(1,0) + y(0,1),$$

so, from Theorem C39,

$$\begin{aligned} r_\theta(x,y) &= r_\theta(x(1,0) + y(0,1)) \\ &= x r_\theta(1,0) + y r_\theta(0,1) \\ &= x(\cos\theta, \sin\theta) + y(-\sin\theta, \cos\theta) \\ &= (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta). \end{aligned}$$

This method of finding an algebraic definition for r_θ is simpler than the geometric approach used in Subsection 1.1 and is more generally applicable.

Exercise C88

Find the image of a vector (x,y) in $\mathbb{R}^2$ under the reflection q_ϕ, given that $q_\phi(1,0) = (\cos 2\phi, \sin 2\phi)$ and $q_\phi(0,1) = (\sin 2\phi, -\cos 2\phi)$.

2 Matrices of linear transformations

In this section you will see how the images of basis vectors can be used to find the matrix representation of a linear transformation.

2.1 Finding matrix representations

In Section 1 you met several examples of matrix representations of linear transformations. For example, you saw that a k-dilation of $\mathbb{R}^2$ can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix}$$

and a rotation r_θ of $\mathbb{R}^2$ can be represented by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

In this section we show that any linear transformation $t : V \longrightarrow W$ between finite-dimensional vector spaces has a matrix representation

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \longmapsto \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix},$$

or

$$\mathbf{v}^T \longmapsto \mathbf{A}\mathbf{v}^T = \mathbf{w}^T.$$

Matrix representations are important because they are an aid to performing calculations with linear transformations; in particular, they are easily handled by computers.

You have seen that it is sometimes convenient to use a non-standard basis $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for a vector space V. Recall from Unit C2 that if $\mathbf{v}$ is a vector in V and

$$\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n,$$

then the numbers $v_1, \ldots, v_n$ are the *coordinates* of $\mathbf{v}$ with respect to the basis E (the *E-coordinates* of $\mathbf{v}$). The *E-coordinate representation* of $\mathbf{v}$ is $\mathbf{v}_E = (v_1, \ldots, v_n)_E$.

For example, let E be the basis $\{(1, 1), (1, 0)\}$ for $\mathbb{R}^2$. The vector $\mathbf{v} = (5, 2)$ in $\mathbb{R}^2$ can be written as

$$\mathbf{v} = 2(1, 1) + 3(1, 0),$$

so the E-coordinate representation of $\mathbf{v}$ is

$$\mathbf{v}_E = (2, 3)_E.$$

For another example, consider the basis $E = \{1 + x^2, x^2, 2 - x\}$ for the vector space P_3. As

$$1 + x + 2x^2 = 3(1 + x^2) - x^2 - (2 - x),$$

the E-coordinate representation of the polynomial $1 + x + 2x^2$ is $(3, -1, -1)_E$.

The following exercises should remind you how to write a vector in terms of its coordinates with respect to a given basis.

Exercise C89

Find the E-coordinate representation of the vector $\mathbf{v} = (3, 1)$ in $\mathbb{R}^2$ for each of the following bases E for $\mathbb{R}^2$.

(a) $E = \{(3, 1), (2, 1)\}$ (b) $E = \{(1, 2), (2, 1)\}$

Exercise C90

Find the E-coordinate representation of the polynomial $p(x) = 2 + 3x$ in P_2 for each of the following bases E for P_2.

(a) $E = \{1, x\}$ (the standard basis)

(b) $E = \{1, 4 + 6x\}$

(c) $E = \{2x, 1 + 4x\}$

We now define a *matrix representation* of a linear transformation between finite-dimensional vector spaces, with respect to specified bases.

Definition

Let V and W be vector spaces of dimensions n and m, respectively. Let $t : V \longrightarrow W$ be a linear transformation, let $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for V, let $F = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ be a basis for W and let $\mathbf{A}$ be an $m \times n$ matrix such that

$$t(\mathbf{v})_F = \mathbf{A}\mathbf{v}_E, \quad \text{for each vector } \mathbf{v} \text{ in } V.$$

Then $\mathbf{v}_E \longmapsto \mathbf{A}\mathbf{v}_E = t(\mathbf{v})_F$ is the **matrix representation** of t with respect to the bases E and F, and $\mathbf{A}$ is the **matrix** of t with respect to the bases E and F.

Remarks

1. A matrix of a linear transformation from an n-dimensional vector space to an m-dimensional vector space is an $m \times n$ matrix, not an $n \times m$ matrix as you might expect.
2. Strictly speaking, since we defined vectors as *row vectors*, we should write $\mathbf{v}_E^T \longmapsto \mathbf{A}\mathbf{v}_E^T = t(\mathbf{v})_F^T$. However, we omit the transpose symbols for simplicity, and we often write these vectors as row vectors to save space.
3. When $E = F$, we refer to the matrix representation with respect to the basis E.

Later in this section we will *prove* that there is exactly one matrix of t with respect to the bases E and F, but first we develop a strategy (Strategy C15) for finding the matrix of a linear transformation.

Matrix representations using standard bases

We start by considering linear transformations where both E and F are the standard basis.

Exercise C91

Each of the following linear transformations $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is defined by a matrix representation with respect to the standard basis $\{(1,0),(0,1)\}$ for $\mathbb{R}^2$. In each case, find the images of the vectors $(1,0)$ and $(0,1)$. What do you notice about the relationship between the vectors $t(1,0)$ and $t(0,1)$ and the 2×2 matrix of the linear transformation?

(a) A $(3,2)$-scaling of $\mathbb{R}^2$.

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

(b) A rotation $r_{\pi/4}$ of $\mathbb{R}^2$.

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

In Exercise C91 you saw two examples in which, given a transformation defined by a matrix, the coordinates of the images of the standard basis vectors of the domain were the columns of the matrix. It turns out that this is always the case, even for non-standard bases: the coordinates of the images of the basis vectors of the domain are the columns of the matrix. This gives a strategy for finding the matrix of a linear transformation between any two finite-dimensional vector spaces with respect to *any* bases for the domain and codomain.

Strategy C15

To find the matrix $\mathbf{A}$ of a linear transformation $t : V \longrightarrow W$ with respect to the basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for V, and the basis $F = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ for W, do the following.

1. Find $t(\mathbf{e}_1), t(\mathbf{e}_2), \ldots, t(\mathbf{e}_n)$.
2. Find the F-coordinates of each of these image vectors.

$$\begin{aligned} t(\mathbf{e}_1) &= (\mathbf{a}_{11}, \mathbf{a}_{21}, \ldots, \mathbf{a}_{m1})_F \\ t(\mathbf{e}_2) &= (\mathbf{a}_{12}, \mathbf{a}_{22}, \ldots, \mathbf{a}_{m2})_F \\ &\;\;\vdots \\ t(\mathbf{e}_n) &= (\mathbf{a}_{1n}, \mathbf{a}_{2n}, \ldots, \mathbf{a}_{mn})_F \end{aligned}$$

3. Construct the matrix $\mathbf{A}$ column by column.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

We first illustrate the strategy with some exercises and then prove that it works later in this section.

Worked Exercise C53

For each of the following linear transformations t, find the matrix representation of t with respect to the standard bases for the domain and codomain.

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$
$(x, y) \longmapsto (2x, 3x + y, y)$

(b) $t : P_3 \longrightarrow P_2$
$p(x) \longmapsto p'(x)$

Solution

(a) We use Strategy C15.

The standard basis for $\mathbb{R}^2$ is $\{(1,0),(0,1)\}$.

We find the images of the vectors in the domain basis $E = \{(1,0),(0,1)\}$:

$$t(1,0) = (2,3,0), \quad t(0,1) = (0,1,1).$$

The standard basis for $\mathbb{R}^3$ is $\{(1,0,0),(0,1,0),(0,0,1)\}$. There is really nothing to do here when the basis in the codomain is the standard basis since the images are already expressed with respect to this basis; we show the working here for completeness.

We find the F-coordinates of each of these image vectors, where $F = \{(1,0,0),(0,1,0),(0,0,1)\}$:

$$t(1,0) = (2,3,0)_F, \quad t(0,1) = (0,1,1)_F.$$

We now construct the matrix of t by writing down the coordinates of the image vectors column by column – keeping the columns in the same order as the corresponding domain basis vectors.

Hence the matrix of t with respect to the standard bases for the domain and codomain is

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 3 & 1 \\ 0 & 1 \end{pmatrix}.$$

Thus the matrix representation of t with respect to these bases is

$$\begin{pmatrix} x \\ y \end{pmatrix}_E \longmapsto \begin{pmatrix} 2 & 0 \\ 3 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}_E = \begin{pmatrix} 2x \\ 3x + y \\ y \end{pmatrix}_F.$$

We have included the subscripts E and F here, but often will omit these where the bases are the standard ones.

(b) The standard basis for P_3 is $\{1, x, x^2\}$. Therefore the three basis vectors and their derivatives are as follows: $p_1(x) = 1$ and $p_1'(x) = 0$, $p_2(x) = x$ and $p_2'(x) = 1$, $p_3(x) = x^2$ and $p_3'(x) = 2x$.

We find the images of the vectors in the domain basis $E = \{1, x, x^2\}$:

$$t(1) = 0, \quad t(x) = 1, \quad t(x^2) = 2x.$$

The standard basis for P_2 is $\{1, x\}$. We notice that $0 = 0 + 0x$, $1 = 1 + 0x$ and $2x = 0 + 2x$.

We find the F-coordinates of each of these image vectors, where $F = \{1, x\}$:

$$t(1) = (0, 0)_F, \quad t(x) = (1, 0)_F, \quad t(x^2) = (0, 2)_F.$$

We keep the columns in this order.

Hence the matrix of t with respect to the standard bases for the domain and codomain is

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Thus the matrix representation of t with respect to these bases for P_3 and P_2 is

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}_E \longmapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}_E = \begin{pmatrix} b \\ 2c \end{pmatrix}_F.$$

We have $t(a + bx + cx^2) = b + 2cx$.

Exercise C92

For each of the following linear transformations t, find the matrix representation of t with respect to the standard bases for the domain and codomain.

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x + 3y, y)$

(b) $t : P_3 \longrightarrow P_3$
$p(x) \longmapsto p(x) + p(2)$

(c) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$
$(x, y) \longmapsto (x, y, x, y)$

(d) $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$
$(x, y, z) \longmapsto (x, y)$

Matrix representations using non-standard bases

So far we have used the strategy to find matrix representations with respect to standard bases. We now use the strategy to find matrix representations with respect to other bases.

We start with a non-standard basis for the domain and the standard basis for the codomain.

Worked Exercise C54

Find the matrix representation of the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x, y) \end{aligned}$$

with respect to the non-standard domain basis $E = \{(1,1,1), (1,1,0), (1,0,0)\}$ and the standard codomain basis $F = \{(1,0), (0,1)\}$.

Solution

We use Strategy C15.

We find the images of the vectors in the domain basis $E = \{(1,1,1), (1,1,0), (1,0,0)\}$:

$$t(1,1,1) = (1,1), \quad t(1,1,0) = (1,1), \quad t(1,0,0) = (1,0).$$

We find the F-coordinates of each of these image vectors, where $F = \{(1,0), (0,1)\}$:

$$t(1,1,1) = (1,1)_F, \quad t(1,1,0) = (1,1)_F, \quad t(1,0,0) = (1,0)_F.$$

We keep the columns in this order.

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Thus the matrix representation of t with respect to the non-standard basis E for $\mathbb{R}^3$ and the standard basis F for $\mathbb{R}^2$ is

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}_E \longmapsto \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}_E = \begin{pmatrix} v_1 + v_2 + v_3 \\ v_1 + v_2 \end{pmatrix}_F$$

Using v_1, v_2 and v_3 instead of x, y and z helps emphasise that these are coordinates with respect to a non-standard basis.

Compare the matrix representation in Worked Example C54 to that found in Exercise C92(d) for this linear transformation with respect to the

standard basis in both the domain and the codomain:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

In general, different bases give different matrix representations.

We now consider a non-standard basis in the codomain while keeping to the standard basis in the domain.

Worked Exercise C55

Find the matrix representation of the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^4 \\ (x, y) &\longmapsto (x, y, x, y) \end{aligned}$$

with respect to the standard domain basis $E = \{(1,0),(0,1)\}$ and the non-standard codomain basis $F = \{(1,0,0,0),(1,1,0,0),(1,1,1,0),(1,1,1,1)\}$.

Solution

We use Strategy C15.

We find the images of the vectors in the domain basis $E = \{(1,0),(0,1)\}$:

$$t(1,0) = (1,0,1,0), \quad t(0,1) = (0,1,0,1).$$

We now write these image vectors in terms of their coordinates with respect to the codomain basis – this requires some work!

We find the F-coordinates of each of these image vectors, where $F = \{(1,0,0,0),(1,1,0,0),(1,1,1,0),(1,1,1,1)\}$.

For the first image vector, we need $a, b, c, d \in \mathbb{R}$ such that

$$(1,0,1,0) = (a,b,c,d)_F.$$

Since

$$\begin{aligned} (a,b,c,d)_F &= a(1,0,0,0) + b(1,1,0,0) + c(1,1,1,0) + d(1,1,1,1) \\ &= (a+b+c+d, b+c+d, c+d, d), \end{aligned}$$

by equating coordinates we obtain the following system

$$\begin{aligned} a + b + c + d &= 1 \\ b + c + d &= 0 \\ c + d &= 1 \\ d &= 0. \end{aligned}$$

Solving, we have $d = 0$, $c = 1$, $b = -1$ and $a = 1$, so

$$(1,0,1,0) = (1,-1,1,0)_F.$$

Therefore

$$t(1,0) = (1,-1,1,0)_F.$$

We have found the F-coordinates of the image of the first domain basis vector. If the equations had been more difficult to solve we could have used Gauss–Jordan elimination as we did in Unit C1.

For the second image vector, we need $e, f, g, h \in \mathbb{R}$ such that

$$(0,1,0,1) = (e,f,g,h)_F.$$

Since

$$\begin{aligned}(e,f,g,h)_F &= e(1,0,0,0) + f(1,1,0,0) + g(1,1,1,0) + h(1,1,1,1)\\ &= (e+f+g+h, f+g+h, g+h, h),\end{aligned}$$

by equating coordinates we obtain the system

$$\begin{aligned} e + f + g + h &= 0\\ f + g + h &= 1\\ g + h &= 0\\ h &= 1.\end{aligned}$$

Solving, we have $h = 1$, $g = -1$, $f = 1$ and $e = -1$, so

$$(0,1,0,1) = (-1,1,-1,1)_F.$$

Therefore

$$t(0,1) = (-1,1,-1,1)_F.$$

We keep the columns in this order.

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} 1 & -1\\ -1 & 1\\ 1 & -1\\ 0 & 1\end{pmatrix}.$$

Thus the matrix representation of t with respect to the standard basis E for $\mathbb{R}^2$ and the non-standard basis F for $\mathbb{R}^4$ is

$$\begin{pmatrix} v_1\\ v_2\end{pmatrix}_E \longmapsto \begin{pmatrix} 1 & -1\\ -1 & 1\\ 1 & -1\\ 0 & 1\end{pmatrix}\begin{pmatrix} v_1\\ v_2\end{pmatrix}_E = \begin{pmatrix} v_1 - v_2\\ -v_1 + v_2\\ v_1 - v_2\\ v_2\end{pmatrix}_F.$$

Compare the matrix representation in Worked Exercise C55 to that found in Exercise C92(c) for this linear transformation with respect to the

standard bases in both the domain and codomain:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ x \\ y \end{pmatrix}.$$

Finally, we look at an example with non-standard bases for both the domain and the codomain.

Worked Exercise C56

Find the matrix representation of the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longmapsto (2x, 3x + y, y) \end{aligned}$$

with respect to the non-standard domain basis $E = \{(1, 1), (1, 0)\}$ and the non-standard codomain basis $F = \{(1, 1, 1), (0, 1, 1), (0, 0, 1))\}$.

Solution

We use Strategy C15.

We find the images of the domain basis vectors $E = \{(1, 1), (1, 0)\}$:

$$t(1, 1) = (2, 4, 1), \quad t(1, 0) = (2, 3, 0).$$

We find the F-coordinates of each of these image vectors, where $F = \{(1, 1, 1), (0, 1, 1), (0, 0, 1))\}$.

For the first image vector we need $a, b, c \in \mathbb{R}$ such that

$$(2, 4, 1) = (a, b, c)_F.$$

Since

$$\begin{aligned} (a, b, c)_F &= a(1, 1, 1) + b(0, 1, 1) + c(0, 0, 1) \\ &= (a, a + b, a + b + c), \end{aligned}$$

by equating coordinates we obtain the system

$$\begin{aligned} a \qquad\quad &= 2 \\ a + b \qquad &= 4 \\ a + b + c &= 1. \end{aligned}$$

Solving, we have $a = 2$, $b = 2$ and $c = -3$, so

$$(2, 4, 1) = (2, 2, -3)_F.$$

Therefore

$$t(1, 1) = (2, 2, -3)_F.$$

For the second image vector we need $d, e, f \in \mathbb{R}$ such that

$$(2, 3, 0) = (d, e, f)_F.$$

Since

$$\begin{aligned}(d,e,f)_F &= d(1,1,1)+e(0,1,1)+f(0,0,1)\\ &= (d,d+e,d+e+f),\end{aligned}$$

by equating coordinates we obtain the following system

$$\begin{aligned} d \qquad\quad &= 2\\ d+e \qquad &= 3\\ d+e+f &= 0.\end{aligned}$$

Solving, we have $d=2$, $e=1$ and $f=-3$, so

$$(2,3,0)=(2,1,-3)_F.$$

Therefore

$$t(1,0)=(2,1,-3)_F.$$

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A}=\begin{pmatrix}2 & 2\\ 2 & 1\\ -3 & -3\end{pmatrix}.$$

Thus the matrix representation of t with respect to the non-standard basis E for $\mathbb{R}^2$ and the non-standard basis F for $\mathbb{R}^3$ is

$$\begin{pmatrix}v_1\\ v_2\end{pmatrix}_E \longmapsto \begin{pmatrix}2 & 2\\ 2 & 1\\ -3 & -3\end{pmatrix}\begin{pmatrix}v_1\\ v_2\end{pmatrix}_E = \begin{pmatrix}2v_1+2v_2\\ 2v_1+v_2\\ -3v_1-3v_2\end{pmatrix}_F.$$

Compare the matrix representation in Worked Exercise C56 to that found in Worked Exercise C53(a) for this linear transformation with respect to the standard bases in both the domain and codomain:

$$\begin{pmatrix}x\\ y\end{pmatrix} \longmapsto \begin{pmatrix}2 & 0\\ 3 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}2x\\ 3x+y\\ y\end{pmatrix}.$$

The following exercise involves both standard and non-standard bases in the domain and codomain.

Exercise C93

Find the matrix of the linear transformation

$$\begin{aligned} t:\mathbb{R}^3 &\longrightarrow \mathbb{R}^2\\ (x,y,z) &\longmapsto (x,y)\end{aligned}$$

with respect to each of the following bases E for $\mathbb{R}^3$ and F for $\mathbb{R}^2$.

(a) $E=\{(1,0,1),(1,0,0),(1,1,1)\}$
$F=\{(1,0),(0,1)\}$ (standard basis for $\mathbb{R}^2$)

(b) $E = \{(1,0,0),(0,1,0),(0,0,1)\}$ (standard basis for $\mathbb{R}^3$)
$F = \{(2,1),(1,1)\}$

(c) $E = \{(0,1,0),(1,1,1),(0,1,1)\}$ $\quad F = \{(1,3),(2,4)\}$

You have seen that a linear transformation $t : V \longrightarrow W$ has different matrix representations depending on the bases used for the domain and codomain. Moreover, the order of the elements in a basis is important. For example, in the next exercise you should obtain different matrices for t for each part: although the bases contain the same elements, the order in which they appear in the domain basis is different.

In summary, note the following two facts.

- *Different bases* for V and W give *different* matrix representations.
- A *different order* of basis elements gives a *different* matrix representation.

Exercise C94

Find the matrix representation of the linear transformation

$$\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x) \end{aligned}$$

with respect to each of the following bases E for P_3 and F for P_2.

(a) $E = \{1, x, x^2\}$ $\quad F = \{2x, 1 + x\}$

(b) $E = \{x, x^2, 1\}$ $\quad F = \{2x, 1 + x\}$

The unique matrix representation of a linear transformation

You have seen that the matrix representation of a linear transformation depends on the bases for both the domain and the codomain, and the order of these basis elements. However, for given ordered basis elements, there is precisely *one* matrix representation: the one given by Strategy C15. Using the notation in the strategy, the unique matrix representation of a linear transformation t with respect to the bases E and F is

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_E \longmapsto \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_E = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ a_{21}v_1 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix}_F.$$

We now prove this result. If you are short of time, you should skim through this proof and come back to it when time permits.

Theorem C40

Let $t : V \longrightarrow W$ be a linear transformation, let $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for V and let $F = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ be a basis for W. Let

$$\begin{aligned}
t(\mathbf{e}_1) &= (a_{11}, a_{21}, \ldots, a_{m1})_F,\\
t(\mathbf{e}_2) &= (a_{12}, a_{22}, \ldots, a_{m2})_F,\\
&\ \ \vdots\\
t(\mathbf{e}_n) &= (a_{1n}, a_{2n}, \ldots, a_{mn})_F.
\end{aligned}$$

Then there is exactly one matrix of t with respect to the bases E and F, namely

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Proof We start by showing that $\mathbf{A}$ is a matrix of t with respect to the (ordered) bases E and F.

Suppose that the conditions of the theorem are satisfied and that $(v_1, \ldots, v_n)_E$ is the E-coordinate representation of a vector $\mathbf{v} \in V$. Then we have

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n.$$

By Theorem C39, linear transformations preserve linear combinations of vectors, so

$$\begin{aligned}
t(\mathbf{v}) &= t(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n)\\
&= v_1 t(\mathbf{e}_1) + v_2 t(\mathbf{e}_2) + \cdots + v_n t(\mathbf{e}_n)\\
&= v_1(a_{11}, a_{21}, \ldots, a_{m1})_F + v_2(a_{12}, a_{22}, \ldots, a_{m2})_F + \cdots\\
&\quad + v_n(a_{1n}, a_{2n}, \ldots, a_{mn})_F\\
&= (v_1a_{11} + \cdots + v_na_{1n},\ v_1a_{21} + \cdots + v_na_{2n},\ \ldots,\\
&\qquad v_1a_{m1} + \cdots + v_na_{mn})_F.
\end{aligned}$$

So the first coordinate of $t(\mathbf{v})$ is $a_{11}v_1 + \cdots + a_{1n}v_n$, the second coordinate of $t(\mathbf{v})$ is $a_{21}v_1 + \cdots + a_{2n}v_n$, and so on. These coordinates can be obtained by matrix multiplication as follows

$$\begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ a_{21}v_1 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Therefore

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_E \longmapsto \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_E = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ a_{21}v_1 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix}_F$$

is a matrix representation of t with respect to the bases E and F, and

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is a matrix of t with respect to the bases E and F.

We now show that $\mathbf{A}$ is the *only* possible matrix of t with respect to the (ordered) bases E and F. We do this by assuming that there is another possible matrix $\mathbf{B}$ and concluding that $\mathbf{B}$ must be equal to $\mathbf{A}$.

Suppose that $\mathbf{B}$ is also a matrix of t with respect to the bases E and F where

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}.$$

Since $\mathbf{e}_1$ is the first basis vector in E, we have $\mathbf{e}_1 = (1, 0, \ldots, 0)_E$, and the image of $\mathbf{e}_1$ under t is

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_E \longmapsto \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_E = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}_E;$$

that is,

$$t(\mathbf{e}_1) = (b_{11}, b_{21}, \ldots, b_{m1})_F.$$

However,

$$t(\mathbf{e}_1) = (a_{11}, a_{21}, \ldots, a_{m1})_F,$$

so the first column of $\mathbf{B}$ is equal to the first column of $\mathbf{A}$.

Similarly, we find that

$$\begin{aligned} t(\mathbf{e}_2) &= (b_{12}, b_{22}, \ldots, b_{m2})_F = (a_{12}, a_{22}, \ldots, a_{m2})_F, \\ t(\mathbf{e}_3) &= (b_{13}, b_{23}, \ldots, b_{m3})_F = (a_{13}, a_{23}, \ldots, a_{m3})_F, \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ t(\mathbf{e}_n) &= (b_{1n}, b_{2n}, \ldots, b_{mn})_F = (a_{1n}, a_{2n}, \ldots, a_{mn})_F. \end{aligned}$$

Therefore each subsequent column of $\mathbf{B}$ is also the same as the corresponding column of $\mathbf{A}$. Since $\mathbf{A}$ and $\mathbf{B}$ are both $m \times n$ matrices and their corresponding entries are equal, we have $\mathbf{B} = \mathbf{A}$.

Thus $\mathbf{A}$ is the only matrix of t with respect to the bases E and F. ■

2.2 An equivalent definition

We have shown that any linear transformation $t : V \longrightarrow W$, where V and W are finite-dimensional vector spaces, has a matrix representation. We now show the converse – that a function that has a matrix representation is a linear transformation. We will use the following result about matrix multiplication: if $\mathbf{A}$ and $\mathbf{B}$ are matrices and α a scalar, then $(\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$, whenever this product exists. You might like to prove this result yourself; it is included as a 'challenging' exercise in the additional exercises booklet for this unit.

Theorem C41

Let $t : V \longrightarrow W$ be a function that has a matrix representation. Then t is a linear transformation.

Proof Suppose that the function $t : V \longrightarrow W$ has a matrix representation

$$\mathbf{v}_E \longmapsto \mathbf{A}\mathbf{v}_E = t(\mathbf{v})_F.$$

We first show that t satisfies LT1: that for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ we have $t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2)$.

Let $\mathbf{v}_1, \mathbf{v}_2 \in V$. Then

$$t(\mathbf{v}_1 + \mathbf{v}_2)_F = \mathbf{A}(\mathbf{v}_1 + \mathbf{v}_2)_E$$

and

$$\begin{aligned} t(\mathbf{v}_1)_F + t(\mathbf{v}_2)_F &= \mathbf{A}(\mathbf{v}_1)_E + \mathbf{A}(\mathbf{v}_2)_E \\ &= \mathbf{A}(\mathbf{v}_1 + \mathbf{v}_2)_E, \end{aligned}$$

by the distributive property for matrix multiplication.

So $t(\mathbf{v}_1 + \mathbf{v}_2)_F = t(\mathbf{v}_1)_F + t(\mathbf{v}_2)_F$, and hence $t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2)$, because the F-coordinate representation of a vector is unique. Therefore the function t satisfies LT1.

We now show that t satisfies LT2: that for all $\mathbf{v} \in V$ and $\alpha \in \mathbb{R}$ we have $t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v})$.

Let $\mathbf{v} \in V$ and $\alpha \in \mathbb{R}$. Then

$$\alpha\, t(\mathbf{v})_F = \alpha\mathbf{A}\mathbf{v}_E$$

and

$$t(\alpha\mathbf{v})_F = \mathbf{A}(\alpha\mathbf{v})_E = \alpha\mathbf{A}\mathbf{v}_E,$$

by the result about matrix multiplication quoted above.

So $t(\alpha\mathbf{v})_F = \alpha\, t(\mathbf{v})_F$, and hence $t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v})$, because the F-coordinate representation of a vector is unique. Therefore the function t also satisfies LT2.

Since both LT1 and LT2 are satisfied, the function t is a linear transformation. ∎

Theorems C40 and C41 imply the following.

Corollary C42

A function $t : V \longrightarrow W$, where V and W are finite-dimensional vector spaces, is a linear transformation if and only if it has a matrix representation.

This means, for example, that the linear transformations from $\mathbb{R}^2$ to itself are those functions that have a matrix representation

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

So the linear transformations from $\mathbb{R}^2$ to itself are those functions of the form

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (ax + by, cx + dy) \end{aligned} \tag{1}$$

for some $a, b, c, d \in \mathbb{R}$.

Similar expressions exist for linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Exercise C95

Use the linear transformation form (1) to determine which of the following functions are linear transformations.

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (y, 2x + y)$

(b) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x^2, y)$

(c) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x, 2xy + y)$

(d) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (3x, x + 4y)$

3 Composition and invertibility

In this section you will use the matrix representation of a linear transformation to find composite linear transformations and investigate properties of linear transformations, such as invertibility.

3.1 Composition Rule

In the previous section you saw that a function $t : V \longrightarrow W$, where V and W are finite-dimensional vector spaces, is a linear transformation if and only if it has a matrix representation. We now use some of the properties of matrices that you met in Unit C1 to develop our understanding of linear transformations.

We begin by considering the composition of linear transformations. The composite of two functions $t : V \longrightarrow W$ and $s : W \longrightarrow X$ is

$$\begin{aligned} s \circ t : V &\longrightarrow X \\ \mathbf{v} &\longmapsto s(t(\mathbf{v})), \end{aligned}$$

as shown in Figure 18.

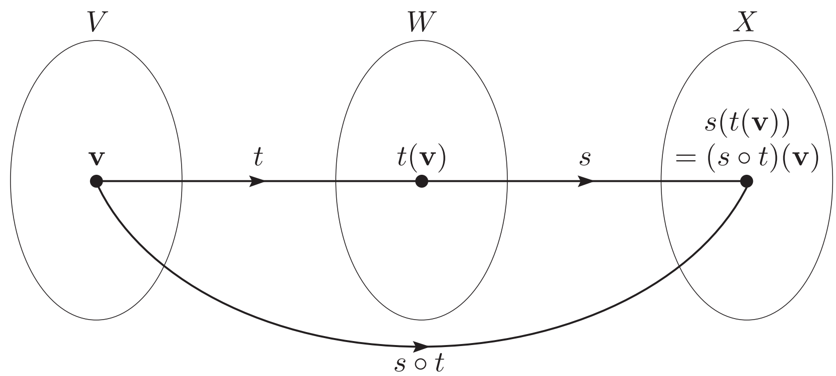


Figure 18 The composite $s \circ t$

Consider the linear transformations

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x + 2y, y) \end{aligned} \quad \text{and} \quad \begin{aligned} s : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (5x, x + y). \end{aligned} \tag{2}$$

Let $(x, y) \in \mathbb{R}^2$. Then

$$t(x, y) = (x + 2y, y),$$

so

$$\begin{aligned} s(t(x, y)) &= s(x + 2y, y) \\ &= (5(x + 2y), (x + 2y) + y) \\ &= (5x + 10y, x + 3y). \end{aligned}$$

Thus the composite function $s \circ t$ is the linear transformation

$$\begin{aligned} s \circ t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (5x + 10y, x + 3y). \end{aligned}$$

In general, for linear transformations s and t from a vector space to itself, the composite functions $s \circ t$ and $t \circ s$ are not the same, as you will see in the following exercise.

Exercise C96

Let p and r be the linear transformations

$$\begin{aligned} p : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (3x + y, -x) \end{aligned} \quad \text{and} \quad \begin{aligned} r : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, x + y). \end{aligned}$$

Find the following composite functions.

(a) $r \circ p$ (b) $p \circ r$

Each of the composite functions in Exercise C96 is a linear transformation, since it has the correct form (1). In the next theorem (Theorem C43) we show that composition of two linear transformations always gives a linear transformation.

At the beginning of this subsection we showed that the two linear transformations s and t in equation (2) can be composed to give the linear transformation

$$\begin{aligned} s \circ t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (5x + 10y, x + 3y). \end{aligned}$$

Using Strategy C15, we obtain the matrix representations of these three linear transformations with respect to the standard basis for $\mathbb{R}^2$:

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\longmapsto \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ y \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} s : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\longmapsto \begin{pmatrix} 5 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x \\ x + y \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} s \circ t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\longmapsto \begin{pmatrix} 5 & 10 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x + 10y \\ x + 3y \end{pmatrix}. \end{aligned}$$

We can check that

$$\begin{pmatrix} 5 & 10 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},$$

so, in this example,

$$\begin{pmatrix} \text{matrix} \\ \text{of } s \circ t \end{pmatrix} = \begin{pmatrix} \text{matrix} \\ \text{of } s \end{pmatrix} \begin{pmatrix} \text{matrix} \\ \text{of } t \end{pmatrix}.$$

We now show that this relationship between the matrices of $s \circ t$, s and t holds in general; that is, that composition of linear transformations corresponds to matrix multiplication. If you are short of time, you should just look at the structure of this proof and come back to it when time permits; part (a) checks the properties LT1 and LT2, and part (b) constructs the matrix of the composite linear transformation. To help visualise what is going on, the composite $s \circ t$, vector spaces and bases are shown in Figure 19.

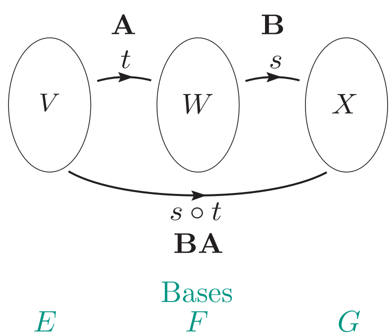


Figure 19 The composite $s \circ t$ showing the vector spaces and bases

Theorem C43 Composition Rule

Let $t : V \longrightarrow W$ and $s : W \longrightarrow X$ be linear transformations. Then:

(a) $s \circ t : V \longrightarrow X$ is a linear transformation

(b) if $\mathbf{A}$ is the matrix of t with respect to the bases E and F, and $\mathbf{B}$ is the matrix of s with respect to the bases F and G, then $\mathbf{BA}$ is the matrix of $s \circ t$ with respect to the bases E and G.

Proof Let $t : V \longrightarrow W$ and $s : W \longrightarrow X$ be linear transformations.

(a) We first show that $s \circ t$ satisfies LT1: that for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ we have $(s \circ t)(\mathbf{v}_1 + \mathbf{v}_2) = (s \circ t)(\mathbf{v}_1) + (s \circ t)(\mathbf{v}_2)$.

Let $\mathbf{v}_1, \mathbf{v}_2 \in V$. Then, since t and s both satisfy LT1, we have

$$\begin{aligned}(s \circ t)(\mathbf{v}_1 + \mathbf{v}_2) &= s(t(\mathbf{v}_1 + \mathbf{v}_2)) \\ &= s(t(\mathbf{v}_1) + t(\mathbf{v}_2)) \\ &= s(t(\mathbf{v}_1)) + s(t(\mathbf{v}_2)).\end{aligned}$$

We also have

$$(s \circ t)(\mathbf{v}_1) + (s \circ t)(\mathbf{v}_2) = s(t(\mathbf{v}_1)) + s(t(\mathbf{v}_2)).$$

So $(s \circ t)(\mathbf{v}_1 + \mathbf{v}_2) = (s \circ t)(\mathbf{v}_1) + (s \circ t)(\mathbf{v}_2)$. Therefore the composite $s \circ t$ satisfies LT1.

We now show that $s \circ t$ satisfies LT2: that for all $\mathbf{v} \in V$, $\alpha \in \mathbb{R}$ we have $(s \circ t)(\alpha\mathbf{v}) = \alpha(s \circ t)(\mathbf{v})$.

Let $\mathbf{v} \in V$ and $\alpha \in \mathbb{R}$. Then, since t and s both satisfy LT2, we have

$$(s \circ t)(\alpha\mathbf{v}) = s(t(\alpha\mathbf{v})) = s(\alpha\, t(\mathbf{v})) = \alpha\, s(t(\mathbf{v})).$$

We also have

$$\alpha(s \circ t)(\mathbf{v}) = \alpha\, s(t(\mathbf{v})).$$

So $(s \circ t)(\alpha\mathbf{v}) = \alpha(s \circ t)(\mathbf{v})$. Therefore the composite $s \circ t$ satisfies LT2.

Since both LT1 and LT2 are satisfied, the composite $s \circ t$ is a linear transformation.

(b) Suppose that the vector spaces V, W and X have dimensions n, m and p, respectively. Then $\mathbf{A}$ is an $m \times n$ matrix of the form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

and $\mathbf{B}$ is a $p \times m$ matrix of the form

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pm} \end{pmatrix}.$$

We use Strategy C15 to find the matrix of the linear transformation $s \circ t$ with respect to the bases E and G.

We find the images under $s \circ t$ of the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ that form the basis E for V.

To find the image of the basis vector $\mathbf{e}_1$, we use the $n \times 1$ column matrix containing the coordinates of $\mathbf{e}_1$ with respect to the basis E. This matrix has 1 in the first row and 0 elsewhere. Using the matrix representations of t and s, we find that

$$t : \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_E \longmapsto \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_E = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}_F$$

and

$$s : \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}_F \longmapsto \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pm} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}_F$$

$$= \begin{pmatrix} b_{11}a_{11} + \cdots + b_{1m}a_{m1} \\ b_{21}a_{11} + \cdots + b_{2m}a_{m1} \\ \vdots \\ b_{p1}a_{11} + \cdots + b_{pm}a_{m1} \end{pmatrix}_G .$$

So

$$(s \circ t)(\mathbf{e}_1) = (b_{11}a_{11} + \cdots + b_{1m}a_{m1}, \dots, b_{p1}a_{11} + \cdots + b_{pm}a_{m1})_G.$$

Similarly, we find that, for $k = 2, \dots, n$,

$$(s \circ t)(\mathbf{e}_k) = (b_{11}a_{1k} + \cdots + b_{1m}a_{mk}, \dots, b_{p1}a_{1k} + \cdots + b_{pm}a_{mk})_G.$$

Next, we find the G-coordinates of each of the image vectors, but the image vectors are already in this form.

We now construct the matrix of $s \circ t$, column by column. The first column contains the coordinates of $(s \circ t)(\mathbf{e}_1)$, the second column contains the coordinates of $(s \circ t)(\mathbf{e}_2)$, and so on. Thus the matrix of $s \circ t$ with respect to the bases E and G is

$$\begin{pmatrix} b_{11}a_{11}+\cdots+b_{1m}a_{m1} & \cdots & b_{11}a_{1k}+\cdots+b_{1m}a_{mk} & \cdots & b_{11}a_{1n}+\cdots+b_{1m}a_{mn} \\ \vdots & & \vdots & & \vdots \\ b_{j1}a_{11}+\cdots+b_{jm}a_{m1} & \cdots & b_{j1}a_{1k}+\cdots+b_{jm}a_{mk} & \cdots & b_{j1}a_{1n}+\cdots+b_{jm}a_{mn} \\ \vdots & & \vdots & & \vdots \\ b_{p1}a_{11}+\cdots+b_{pm}a_{m1} & \cdots & b_{p1}a_{1k}+\cdots+b_{pm}a_{mk} & \cdots & b_{p1}a_{1n}+\cdots+b_{pm}a_{mn} \end{pmatrix}.$$

Using the rules for matrix multiplication, we find that the above matrix is the same as the matrix product $\mathbf{BA}$.

Thus $\mathbf{BA}$ is the matrix of $s \circ t$ with respect to the bases E and G. ■

Worked Exercise C57

Use the Composition Rule to find the matrix representation of the linear transformation $s \circ t$ with respect to the standard bases for the domain and codomain.

$$t : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

and

$$s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Solution

Note that the codomain of t is $\mathbb{R}^2$: when you multiply a 2×3 matrix by a 3×1 one, the result is 2×1.

It follows from the Composition Rule that the matrix of $s \circ t$ with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 6 \\ 8 & 7 & 9 \end{pmatrix}.$$

Thus the matrix representation of $s \circ t$ with respect to the standard bases for the domain and codomain is

$$\begin{aligned} s \circ t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\longmapsto \begin{pmatrix} 2 & 3 & 6 \\ 8 & 7 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} 2x + 3y + 6z \\ 8x + 7y + 9z \end{pmatrix}. \end{aligned}$$

To work out the domain and codomain of $s \circ t$, recall that if $t : V \longrightarrow W$ and $s : W \longrightarrow X$ then $s \circ t : V \longrightarrow X$ (see Figure 18).

Exercise C97

Use the Composition Rule to find the matrix representation of the linear transformation $s \circ t$ with respect to the standard bases for the domain and codomain.

$$t : \mathbb{R}^4 \longrightarrow \mathbb{R}^2 \qquad \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 & 2 & 4 \\ 2 & 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

and

$$s : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \qquad \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

We now return to two examples of linear transformations of vector spaces of polynomials:

$$\begin{aligned} t : P_3 &\longrightarrow P_3 \\ p(x) &\longmapsto p(x) + p(2) \end{aligned}$$

and

$$\begin{aligned} s : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x). \end{aligned}$$

We compose these linear transformations as follows:

$$\begin{aligned} (s \circ t)(p(x)) &= s(t(p(x))) \\ &= s(p(x) + p(2)) \\ &= (p(x) + p(2))' \\ &= p'(x). \end{aligned}$$

Thus the composite is

$$\begin{aligned} s \circ t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x). \end{aligned}$$

In this case, the functions $s \circ t$ and s are the same function.

Exercise C98

Use the Composition Rule to find the matrix representation of the linear transformation $s \circ t$ with respect to the standard bases $E = \{1, x, x^2\}$ for P_3 and $F = \{1, x\}$ for P_2, when

$$\begin{aligned} s : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x) \end{aligned} \quad \text{and} \quad \begin{aligned} t : P_3 &\longrightarrow P_3 \\ p(x) &\longmapsto p(x) + p(2). \end{aligned}$$

(In Worked Exercise C53(b) and Exercise C92(b) you found that s and t have the following matrix representations

$$s : P_3 \longrightarrow P_2$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}_E \longmapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}_E = \begin{pmatrix} b \\ 2c \end{pmatrix}_F$$

and

$$t : P_3 \longrightarrow P_3$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}_E \longmapsto \begin{pmatrix} 2 & 2 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}_E = \begin{pmatrix} 2a + 2b + 4c \\ b \\ c \end{pmatrix}_E$$

with respect to the standard bases E and F for P_3 and P_2, respectively.)

In Subsection 3.2 of Unit C1 we claimed that multiplication of matrices is associative. We now prove this result, by using the Composition Rule (Theorem C43).

Corollary C44

Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be matrices of sizes $q \times p$, $p \times m$ and $m \times n$, respectively. Then

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

Proof Let t, s and r be the linear transformations whose matrix representations with respect to the standard bases for the domain and codomain are

$$\begin{aligned} t : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \mathbf{v} &\longmapsto \mathbf{C}\mathbf{v}, \end{aligned} \qquad \begin{aligned} s : \mathbb{R}^m &\longrightarrow \mathbb{R}^p \\ \mathbf{v} &\longmapsto \mathbf{B}\mathbf{v} \end{aligned} \quad \text{and} \quad \begin{aligned} r : \mathbb{R}^p &\longrightarrow \mathbb{R}^q \\ \mathbf{v} &\longmapsto \mathbf{A}\mathbf{v}. \end{aligned}$$

It follows from the Composition Rule that $\mathbf{A}(\mathbf{BC})$ is the matrix of the linear transformation $r \circ (s \circ t)$ and that $(\mathbf{AB})\mathbf{C}$ is the matrix of the linear transformation $(r \circ s) \circ t$, with respect to the standard bases for the domain and codomain. The linear transformations $r \circ (s \circ t)$ and $(r \circ s) \circ t$ are equal, since $(r \circ (s \circ t))(\mathbf{v})$ and $((r \circ s) \circ t)(\mathbf{v})$ both mean $r(s(t(\mathbf{v})))$. It follows that $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$. ■

This result illustrates how we can prove results about matrices by using linear transformations. We can also prove results about linear transformations by using matrices, as we do in the next subsection.

3.2 Invertible linear transformations

In this subsection we introduce the notion of an *invertible linear transformation*. Suppose that $t : V \longrightarrow W$ is a linear transformation that is one-to-one (no two elements of V have the same image) and is also onto (the image set $t(V)$ is the whole of W); that is, each element of W is the image of exactly one element of V. Then t has an inverse function t^{-1} with domain W, such that

$$t^{-1}(t(\mathbf{v})) = \mathbf{v}, \quad \text{for each } \mathbf{v} \in V,$$

and

$$t(t^{-1}(\mathbf{w})) = \mathbf{w}, \quad \text{for each } \mathbf{w} \in W;$$

that is,

$$t^{-1} \circ t = i_V \quad \text{and} \quad t \circ t^{-1} = i_W.$$

We say that t is *invertible*. This is illustrated in Figure 20.

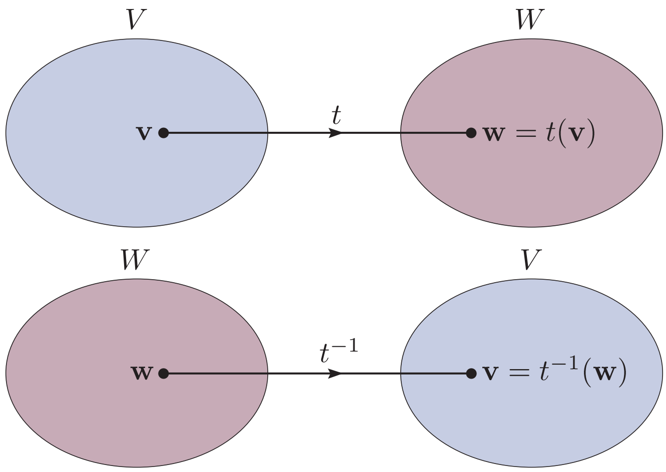


Figure 20 A linear transformation t and its inverse t^{-1}

Definition

The linear transformation $t : V \longrightarrow W$ is **invertible** if there exists an inverse function $t^{-1} : W \longrightarrow V$ such that

$$t^{-1} \circ t = i_V \quad \text{and} \quad t \circ t^{-1} = i_W.$$

Thus a linear transformation $t : V \longrightarrow W$ is invertible if and only if it is one-to-one and onto.

The linear transformation

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (x, 0)$$

is not invertible, since it is not one-to-one; for example,

$$t(1, 1) = t(1, 2) = (1, 0).$$

The linear transformation

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$$
$$(x, y) \longmapsto (x, y, 0)$$

is not invertible, since it is not onto: the image set $t(\mathbb{R}^2)$ is the (x, y)-plane, which is not the whole of the codomain $\mathbb{R}^3$.

Now consider the linear transformation

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (2x, 2y).$$

We can check that t is one-to-one and onto and hence invertible by using the methods of Unit A1, but what is the inverse function of t?

Since t stretches each vector by a factor 2, we expect the inverse function of t to be the linear transformation

$$\begin{aligned} s : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto \left(\tfrac{1}{2}x, \tfrac{1}{2}y\right), \end{aligned}$$

which contracts each vector to half its magnitude. Since

$$s(t(x, y)) = s(2x, 2y) = (x, y)$$

and

$$t(s(x, y)) = t\left(\tfrac{1}{2}x, \tfrac{1}{2}y\right) = (x, y)$$

for each vector (x, y) in $\mathbb{R}^2$, $s \circ t$ and $t \circ s$ are both the identity transformation of $\mathbb{R}^2$, so s is the inverse function of t.

Exercise C99

Verify that the linear transformation

$$\begin{aligned} s : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x + y, 3x + 4y) \end{aligned}$$

is the inverse function of the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (4x - y, -3x + y). \end{aligned}$$

In fact, the inverse of any linear transformation is a linear transformation. Unfortunately, it is not always obvious whether a given linear transformation $t : V \longrightarrow W$ is invertible. Even if we know that t is one-to-one and onto and hence invertible, it may not be clear what the inverse function of t is. If V and W are both *finite*-dimensional vector spaces, however, then t has a matrix representation. The next theorem, illustrated in Figure 21, shows that this can be used to determine whether t is invertible and, if so, to find the inverse function of t. If you are short of time, you should just look at the structure of this proof and come back to it when time permits.

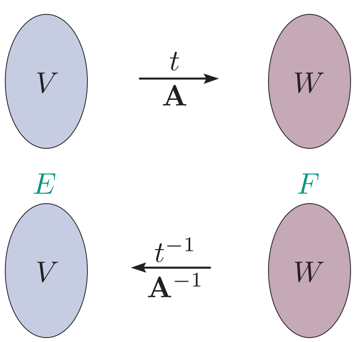


Figure 21 The linear transformation t with matrix $\mathbf{A}$, and its inverse

Theorem C45 Inverse Rule

Let $t : V \longrightarrow W$ be a linear transformation.

(a) If t is invertible, then $t^{-1} : W \longrightarrow V$ is also a linear transformation.

(b) If $\mathbf{A}$ is the matrix of t with respect to the bases E and F, then:

  (i) t is invertible if and only if $\mathbf{A}$ is invertible

  (ii) if t is invertible, then $\mathbf{A}^{-1}$ is the matrix of t^{-1} with respect to the bases F and E.

Proof Let $t : V \longrightarrow W$ be a linear transformation.

(a) Suppose that t is invertible.

We use Strategy C14 to show that the inverse function $t^{-1} : W \longrightarrow V$ is a linear transformation.

We first show that t^{-1} satisfies LT1: for all $\mathbf{w}_1, \mathbf{w}_2 \in W$ we have $t^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = t^{-1}(\mathbf{w}_1) + t^{-1}(\mathbf{w}_2)$.

Let $\mathbf{w}_1, \mathbf{w}_2 \in W$. Then, since t is invertible and hence onto, there exist $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $\mathbf{w}_1 = t(\mathbf{v}_1)$ and $\mathbf{w}_2 = t(\mathbf{v}_2)$. Since t satisfies LT1 we have

$$\begin{aligned} t^{-1}(\mathbf{w}_1 + \mathbf{w}_2) &= t^{-1}(t(\mathbf{v}_1) + t(\mathbf{v}_2)) \\ &= t^{-1}(t(\mathbf{v}_1 + \mathbf{v}_2)) \\ &= \mathbf{v}_1 + \mathbf{v}_2. \end{aligned}$$

Also,

$$\begin{aligned} t^{-1}(\mathbf{w}_1) + t^{-1}(\mathbf{w}_2) &= t^{-1}(t(\mathbf{v}_1)) + t^{-1}(t(\mathbf{v}_2)) \\ &= \mathbf{v}_1 + \mathbf{v}_2. \end{aligned}$$

So $t^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = t^{-1}(\mathbf{w}_1) + t^{-1}(\mathbf{w}_2)$. Therefore t^{-1} satisfies LT1.

Next we show that t^{-1} satisfies LT2: for all $\mathbf{w} \in W$, $\alpha \in \mathbb{R}$ we have $t^{-1}(\alpha\mathbf{w}) = \alpha\, t^{-1}(\mathbf{w})$.

Let $\mathbf{w} \in W$; then there exists $\mathbf{v} \in V$ such that $\mathbf{w} = t(\mathbf{v})$. Let $\alpha \in \mathbb{R}$; then, since t satisfies LT2 we have

$$t^{-1}(\alpha\mathbf{w}) = t^{-1}(\alpha\, t(\mathbf{v})) = t^{-1}(t(\alpha\mathbf{v})) = \alpha\mathbf{v}.$$

Also,

$$\alpha\, t^{-1}(\mathbf{w}) = \alpha\, t^{-1}(t(\mathbf{v})) = \alpha\mathbf{v}.$$

So $t^{-1}(\alpha\mathbf{w}) = \alpha\, t^{-1}(\mathbf{w})$ and LT2 is satisfied.

Since both LT1 and LT2 are satisfied, t^{-1} is a linear transformation.

(b) Let $\mathbf{A}$ be the matrix of t with respect to the bases E and F, so

$$t : \mathbf{v}_E \longmapsto \mathbf{A}\mathbf{v}_E = \mathbf{w}_F, \quad \text{for any vector } \mathbf{v} \in V.$$

We prove the 'if' statement and show that if $\mathbf{A}$ is invertible, then t is invertible. Using properties of matrices, if $\mathbf{A}$ is invertible, then $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$, where $\mathbf{I}$ is the identity matrix.

We show that if $\mathbf{A}$ is invertible, then t is invertible. Suppose that $\mathbf{A}$ is invertible. Then we know that $\mathbf{A}$ is a square matrix and $\mathbf{A}^{-1}$ is also square (and of the same size); so we can define s to be the linear transformation with the matrix representation

$$\begin{aligned} s : W &\longrightarrow V \\ \mathbf{w}_F &\longmapsto \mathbf{A}^{-1}\mathbf{w}_F = s(\mathbf{w})_E. \end{aligned}$$

We show that s is the inverse function of t, and hence that t is invertible.

It follows from the Composition Rule that $s \circ t$ has the matrix representation

$$\begin{aligned} s \circ t : V &\longrightarrow V \\ \mathbf{v}_E &\longmapsto (\mathbf{A}^{-1}\mathbf{A})\mathbf{v}_E = \mathbf{I}\mathbf{v}_E = \mathbf{v}_E. \end{aligned}$$

Thus $s(t(\mathbf{v})) = \mathbf{v}$ for each $\mathbf{v} \in V$; that is, $s \circ t = i_V$.

Similarly, it follows from the Composition Rule that $t \circ s$ has the matrix representation

$$\begin{aligned} t \circ s : W &\longrightarrow W \\ \mathbf{w}_F &\longmapsto (\mathbf{A}\mathbf{A}^{-1})\mathbf{w}_F = \mathbf{I}\mathbf{w}_F = \mathbf{w}_F. \end{aligned}$$

Thus $t(s(\mathbf{w})) = \mathbf{w}$ for each $\mathbf{w} \in W$; that is, $t \circ s = i_W$.

Since $s \circ t = i_V$ and $t \circ s = i_W$, it follows that s is the inverse function of t, so t is invertible.

We prove the 'only if' statement and show that if t is invertible, then $\mathbf{A}$ is invertible *and* $\mathbf{A}^{-1}$ is the matrix of t^{-1} with respect to the bases F and E.

We show that if t is invertible, then $\mathbf{A}$ is invertible. Suppose that t is invertible so t^{-1} is a linear transformation. Then by Theorem C40 it has a matrix representation

$$\begin{aligned} t^{-1} : W &\longrightarrow V \\ \mathbf{w}_F &\longmapsto \mathbf{B}\mathbf{w}_F = t^{-1}(\mathbf{w})_E. \end{aligned}$$

We show that $\mathbf{B} = \mathbf{A}^{-1}$.

It follows from the Composition Rule that $t^{-1} \circ t$ has the matrix representation

$$\begin{aligned} t^{-1} \circ t : V &\longrightarrow V \\ \mathbf{v}_E &\longmapsto (\mathbf{B}\mathbf{A})\mathbf{v}_E. \end{aligned}$$

Since $(t^{-1} \circ t)(\mathbf{v}) = \mathbf{v}$ for each $\mathbf{v} \in V$, it follows that

$$(\mathbf{B}\mathbf{A})\mathbf{v}_E = \mathbf{v}_E, \quad \text{for all } \mathbf{v} \in V.$$

Thus $\mathbf{B}\mathbf{A} = \mathbf{I}$.

Similarly, it follows from the Composition Rule that $t \circ t^{-1}$ has the matrix representation

$$\begin{aligned} t \circ t^{-1} : W &\longrightarrow W \\ \mathbf{w}_F &\longmapsto (\mathbf{A}\mathbf{B})\mathbf{w}_F. \end{aligned}$$

Since $(t \circ t^{-1})(\mathbf{w}) = \mathbf{w}$ for each $\mathbf{w} \in W$, it follows that

$$(\mathbf{A}\mathbf{B})\mathbf{w}_F = \mathbf{w}_F, \quad \text{for all } \mathbf{w} \in W.$$

Thus $\mathbf{A}\mathbf{B} = \mathbf{I}$.

Since

$$\mathbf{B}\mathbf{A} = \mathbf{A}\mathbf{B} = \mathbf{I},$$

it follows that $\mathbf{A}$ is invertible and $\mathbf{B} = \mathbf{A}^{-1}$. Therefore $\mathbf{A}^{-1}$ is the matrix of t^{-1} with respect to the bases F and E.

This completes the proof. ■

One consequence of the Inverse Rule is that if $t : V \longrightarrow W$ is an invertible linear transformation, then any matrix of t must be invertible and hence square. Since a matrix of t has m rows and n columns, where m is the dimension of W and n is the dimension of V, it follows that $m = n$; that is, the vector spaces V and W must have the same dimension, and we have the following corollary to Theorem C45.

Corollary C46

Let $t : V \longrightarrow W$ be an invertible linear transformation, where V and W are finite-dimensional. Then

$$\dim V = \dim W.$$

It follows that if $t : V \longrightarrow W$ is a linear transformation and V and W have *different* finite dimensions, then t is not invertible. For example, the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (2x + y, x - y) \end{aligned}$$

is not invertible, since the domain and codomain have different dimensions.

Now suppose that $t : V \longrightarrow W$ is a linear transformation and that V and W have the *same* finite dimension. It follows from the Inverse Rule that you can use the following strategy to determine whether or not t is invertible. Recall that you saw in Subsection 5.4 of Unit C1 that a matrix is invertible if and only if its determinant is non-zero.

Strategy C16

To determine whether or not a linear transformation $t : V \longrightarrow W$ is invertible, where V and W are n-dimensional vector spaces with bases E and F, respectively, do the following.

1. Find a matrix representation of t,

$$\mathbf{v}_E \longmapsto \mathbf{A}\mathbf{v}_E = t(\mathbf{v})_F.$$

2. Evaluate $\det \mathbf{A}$.

- If $\det \mathbf{A} = 0$, then t is not invertible.
- If $\det \mathbf{A} \neq 0$, then t is invertible and $t^{-1} : W \longrightarrow V$ has the matrix representation

$$\mathbf{w}_F \longmapsto \mathbf{A}^{-1}\mathbf{w}_F = t^{-1}(\mathbf{w})_E.$$

Worked Exercise C58

Show that the following linear transformation t is invertible and find the inverse function of t.

$$t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (x + y, 2y)$$

Solution

We use Strategy C16 and first find a matrix representation of t.

We find a matrix representation of t using Strategy C15.

We have

$$t(1, 0) = (1, 0) \quad \text{and} \quad t(0, 1) = (1, 2).$$

Since we have the standard basis in the codomain, the F-coordinates of the image vectors are immediate.

Hence the matrix representation of t with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ 2y \end{pmatrix}.$$

The next step is to evaluate the determinant of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

We have

$$\det \mathbf{A} = \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = 1 \times 2 - 1 \times 0 = 2.$$

Since $\det \mathbf{A}$ is non-zero, t is invertible.

We now find the inverse function of t. According to Strategy C16, $t^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ has the matrix representation $\mathbf{v} \longmapsto \mathbf{A}^{-1}\mathbf{v}$, with respect to the standard bases for the domain and codomain. Since

$$\mathbf{A}^{-1} = \tfrac{1}{2} \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix},$$

it follows that t^{-1} has the matrix representation

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - \frac{1}{2}y \\ \frac{1}{2}y \end{pmatrix}.$$

So t^{-1} is the linear transformation

$$t^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto \left(x - \tfrac{1}{2}y, \tfrac{1}{2}y\right).$$

Exercise C100

Determine which of the following linear transformations are invertible. Find the inverse function of each invertible linear transformation.

(a) $$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (2x + y, 4x + 2y) \end{aligned}$$

(b) $$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x - y, 3x + y) \end{aligned}$$

(c) $$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (2x, 3y - x, z) \end{aligned}$$

(d) $$\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x) \end{aligned}$$

In Worked Exercise C58 we considered the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x + y, 2y). \end{aligned}$$

We found the matrix $\mathbf{A}$ of t with respect to the *standard* basis for $\mathbb{R}^2$, and showed that $\det \mathbf{A} = 2$. In fact, whatever bases we had chosen for the domain and codomain, we would still have obtained a matrix of t with determinant equal to 2.

It can be shown that the magnitude of the determinant of a matrix of t is the 'scaling factor' of t. Since $\det \mathbf{A} = 2$ in the above case, areas are doubled under t, as shown in Figure 22.

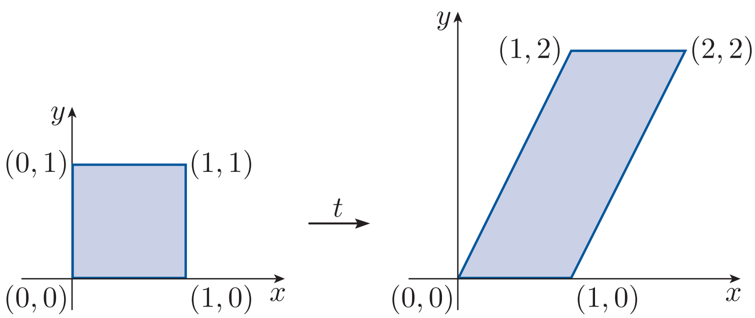


Figure 22 A linear transformation with 'scaling factor' 2

This 'scaling factor' explains the geometric interpretation of the determinant of a 2×2 matrix: that for two position vectors (a, c) and (b, d), the determinant

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

gives the area of the parallelogram with adjacent sides given by these position vectors. The matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is the matrix of the linear transformation with respect to the standard basis for $\mathbb{R}^2$ that maps these basis vectors to (a, c) and (b, d), respectively.

For a linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and a matrix $\mathbf{A}$ of t with $\det \mathbf{A} = 0$, the image of a unit square under t is a line or a point – these have zero area. So, in this case, t is not invertible.

3.3 Isomorphisms

You have seen that there are invertible linear transformations from $\mathbb{R}^2$ to itself, and from $\mathbb{R}^3$ to itself. In fact, whenever the vector spaces V and W have the *same* finite dimension, we can construct an invertible linear transformation from V to W.

For example, consider the two-dimensional vector spaces $\mathbb{R}^2$ and P_2. The linear transformation

$$\begin{aligned} t : P_2 &\longrightarrow \mathbb{R}^2 \\ a + bx &\longmapsto (a, b) \end{aligned}$$

is one-to-one and onto and hence invertible. By looking at a matrix representation of t in this example, we can see how to construct a general invertible linear transformation from V to W, whenever V and W have the same finite dimension.

For t above, take the standard bases $E = \{1, x\}$ for P_2 and $F = \{(1, 0), (0, 1)\}$ for $\mathbb{R}^2$. Then $t(1) = (1, 0)$ and $t(x) = (0, 1)$, so t has the matrix representation

$$\begin{pmatrix} a \\ b \end{pmatrix}_E \longmapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}_E = \begin{pmatrix} a \\ b \end{pmatrix}_F,$$

that is,

$$\mathbf{v}_E \longmapsto \mathbf{I}_2 \mathbf{v}_E = \mathbf{w}_F.$$

More generally, let V and W be n-dimensional vector spaces, let E be a basis for V and let F be a basis for W. Then

$$\begin{aligned} t : V &\longrightarrow W \\ \mathbf{v}_E &\longmapsto \mathbf{I}_n \mathbf{v}_E = \mathbf{w}_F \end{aligned}$$

is a linear transformation from V to W. Since the identity matrix $\mathbf{I}_n$ is invertible, it follows from the Inverse Rule that t is invertible. Note that t maps the first basis vector in E to the first basis vector in F, the second basis vector in E to the second basis vector in F, and so on. We say that t is an *isomorphism* from V to W.

Definition

The vector spaces V and W are **isomorphic** if there exists an invertible linear transformation $t : V \longrightarrow W$. Such a function t is an **isomorphism**.

You met isomorphisms between groups in Unit B2 *Subgroups and isomorphisms*. Isomorphisms between vector spaces are analogous: they identify when vector spaces are 'structurally identical' to each other.

Exercise C101

Write down an isomorphism from P_3 to $\mathbb{R}^3$.

Although the examples of isomorphisms given above involve the identity matrix, any invertible linear transformation provides an isomorphism; so any invertible matrix is possible. For example, consider the following matrix $\mathbf{A}$ of a linear transformation $s : P_3 \longrightarrow \mathbb{R}^3$ with respect to the standard bases in the domain and codomain.

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

This matrix is invertible; you might like to check that it has determinant 2. It is likely that this linear transformation provides a different isomorphism between the vector spaces P_3 and $\mathbb{R}^3$ to the one you wrote down as your answer to Exercise C101.

Suppose that V and W are finite-dimensional vector spaces. We have just seen that if $\dim V = \dim W$, then there is an invertible linear transformation $t : V \longrightarrow W$; that is, V and W are isomorphic.

In particular, each n-dimensional vector space is isomorphic to $\mathbb{R}^n$.

We also know that if V and W have *different* dimensions, then there are *no* invertible linear transformations from V to W; that is, V and W are not isomorphic. Thus we have proved the following result.

Theorem C47

The finite-dimensional vector spaces V and W are isomorphic if and only if

$$\dim V = \dim W.$$

Exercise C102

State which of the following vector spaces are isomorphic to each other:

$$\mathbb{R}^2, \quad \mathbb{R}^3, \quad \mathbb{C}, \quad P_2, \quad P_3.$$

4 Image and kernel

In the previous section you saw that a linear transformation $t : V \longrightarrow W$ is invertible if t is one-to-one and onto; that is, each element of W is the image of exactly one element of V.

In this section you will meet a strategy for determining which elements of W are the images of elements of V, before investigating conditions under which an element of W is the image of more than one element of V. This enables us to prove an important result known as the *Dimension Theorem.*

Finally, we use the Dimension Theorem to show how the number of possible solutions of a system of m linear equations in n unknowns depends on the values of m and n.

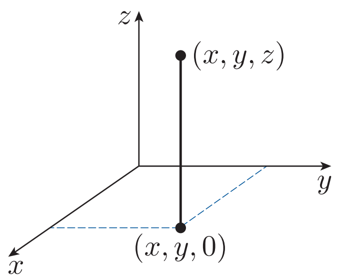


Figure 23 A projection function in $\mathbb{R}^3$

4.1 Image of a linear transformation

Let t be the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x, y, 0). \end{aligned}$$

This projects each vector (x, y, z) onto the vector $(x, y, 0)$ in the (x, y)-plane of $\mathbb{R}^3$ as shown in Figure 23.

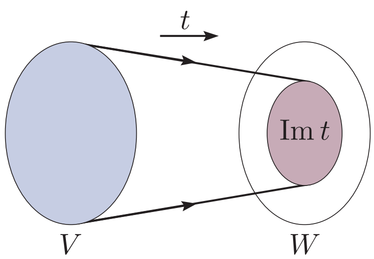


Figure 24 The image set of a linear transformation

A vector $\mathbf{w}$ in the codomain $\mathbb{R}^3$ is the image of a vector $\mathbf{v}$ in the domain $\mathbb{R}^3$ if and only if $\mathbf{w}$ is in the (x, y)-plane. We say that the (x, y)-plane is the *image set* of t. This is a two-dimensional subspace of the codomain $\mathbb{R}^3$.

Recall from Unit A1 that the image set of a function is the set of all elements of the codomain that are images of some element in the domain. Thus the image set of a linear transformation $t : V \longrightarrow W$ is the set of all vectors of W that are images of vectors of V as shown in Figure 24.

The image set of a linear transformation is sometimes simply called its *image.*

Definition

The **image set** of a linear transformation $t : V \longrightarrow W$ is the set

$$\operatorname{Im} t = \{t(\mathbf{v}) : \mathbf{v} \in V\}.$$

Note that the meaning of $\operatorname{Im} t$, which here is the image set of t, is different from that of Im used in Unit A2 *Number systems* where Im meant the imaginary part of a complex number.

It is important to remember that the image set of t is a subset of W, but it need not be equal to W because there may be some vectors of W that are not images of vectors in V. Also, some vectors of W may be images under t of more than one vector of V. Another, equivalent, way of expressing this image set is

$$\operatorname{Im} t = \{\mathbf{w} \in W : \mathbf{w} = t(\mathbf{v}),\ \text{for some } \mathbf{v} \in V\}.$$

Exercise C103

Give a geometric description of the image set of each of the following linear transformations. In each case, state whether the image set is a subspace of the codomain.

(a) $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$
$(x, y, z) \longmapsto (x, 0)$

(b) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x, x)$

For each of the linear transformations in Exercise C103, the image set is a subspace of the codomain. This is true for all linear transformations.

Theorem C48

Let $t : V \longrightarrow W$ be a linear transformation. Then $\operatorname{Im} t$ is a subspace of the codomain W.

Proof We follow Strategy C10 in Unit C2.

We check first that $\mathbf{0} \in \operatorname{Im} t$.

Since t is a linear transformation, $t(\mathbf{0}) = \mathbf{0}$, so $\mathbf{0} \in \operatorname{Im} t$.

This is illustrated in Figure 25.

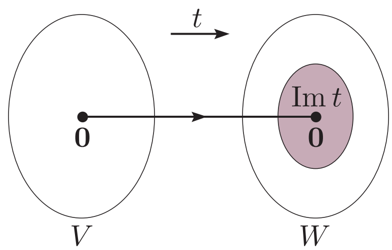


Figure 25 The image of the zero vector

We check next that $\operatorname{Im} t$ is closed under vector addition.

Let $\mathbf{w}_1, \mathbf{w}_2 \in \operatorname{Im} t$. Then there exist vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $\mathbf{w}_1 = t(\mathbf{v}_1)$ and $\mathbf{w}_2 = t(\mathbf{v}_2)$. Since t is a linear transformation,

$$\mathbf{w}_1 + \mathbf{w}_2 = t(\mathbf{v}_1) + t(\mathbf{v}_2) = t(\mathbf{v}_1 + \mathbf{v}_2).$$

Since V is closed under vector addition, $\mathbf{v}_1 + \mathbf{v}_2 \in V$, so $\mathbf{w}_1 + \mathbf{w}_2 \in \operatorname{Im} t$.

This is illustrated in Figure 26.

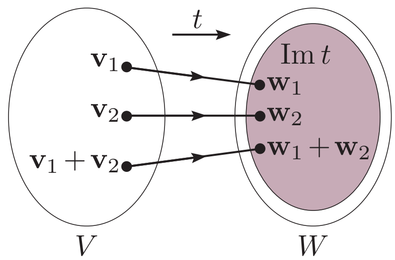


Figure 26 The image of a sum of vectors

Finally, we show that $\operatorname{Im} t$ is closed under scalar multiplication.

Let $\mathbf{w} \in \operatorname{Im} t$ and $\alpha \in \mathbb{R}$. Then there exists $\mathbf{v} \in V$ such that $\mathbf{w} = t(\mathbf{v})$ and, since t is a linear transformation,

$$\alpha \mathbf{w} = \alpha\, t(\mathbf{v}) = t(\alpha \mathbf{v}).$$

Since V is closed under scalar multiplication, $\alpha \mathbf{v} \in V$, so $\alpha \mathbf{w} \in \operatorname{Im} t$.

This is illustrated in Figure 27.

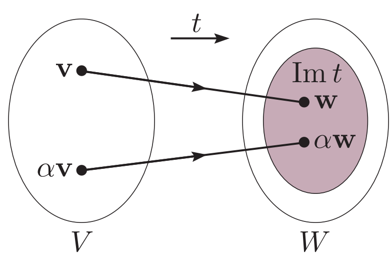


Figure 27 The image of a scalar multiple

Thus $\operatorname{Im} t$ is a subspace of W. ■

For the linear transformations studied so far, it has been easy to write down their image sets. In general, this is not the case; so we need some way of determining the image set of a linear transformation.

If we know the image of each vector in a basis for V, then we can find the image of each vector in V since linear combinations of vectors are preserved (Theorem C39). Thus the image set of t is determined by the images of the domain basis vectors.

For example, consider the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x, y, 0). \end{aligned}$$

The standard basis for the domain $\mathbb{R}^3$ is $\{(1,0,0),(0,1,0),(0,0,1)\}$.

The images of the vectors in this basis are

$$(1,0,0), \quad (0,1,0), \quad (0,0,0).$$

These vectors all lie in $\operatorname{Im} t$, which is the (x, y)-plane, and they *span* $\operatorname{Im} t$; that is, each vector in $\operatorname{Im} t$ can be written as a linear combination of the vectors $(1,0,0)$, $(0,1,0)$ and $(0,0,0)$.

Exercise C104

Let t be the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, x). \end{aligned}$$

Determine the images of the vectors in the standard basis $\{(1,0),(0,1)\}$ for the domain $\mathbb{R}^2$. Do these image vectors span $\operatorname{Im} t$?

(In Exercise C103(b) you found that the image set of this linear transformation is the line $y = x$.)

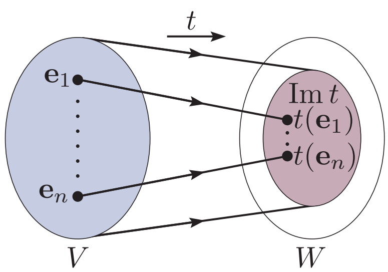


Figure 28 The images of the domain basis vectors in the image set

We now show that, for any linear transformation, the images of the domain basis vectors span the image set; the images of these domain basis vectors are illustrated in Figure 28.

Let $t : V \longrightarrow W$ be a linear transformation and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for V. If $\mathbf{w} \in \operatorname{Im} t$, then $\mathbf{w} = t(\mathbf{v})$ for some $\mathbf{v}$ in V. Since $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for V, there exist real numbers $v_1, \ldots, v_n$ such that

$$\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n.$$

Since t is a linear transformation, it preserves linear combinations of vectors (Theorem C39), so it follows that

$$\mathbf{w} = t(\mathbf{v}) = v_1 t(\mathbf{e}_1) + \cdots + v_n t(\mathbf{e}_n).$$

Thus $\mathbf{w}$ is a linear combination of the vectors $t(\mathbf{e}_1), \ldots, t(\mathbf{e}_n)$.

So $\{t(\mathbf{e}_1), \ldots, t(\mathbf{e}_n)\}$ is a spanning set for $\operatorname{Im} t$, as claimed.

Since a basis is a linearly independent spanning set, we now give a strategy that enables us to find a basis for the image set of a linear transformation.

Strategy C17

To find a basis for $\operatorname{Im} t$, where $t : V \longrightarrow W$ is a linear transformation, do the following.

1. Find a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for the domain V.
2. Determine the vectors $t(\mathbf{e}_1), \ldots, t(\mathbf{e}_n)$.
3. If there is a vector $\mathbf{v}$ in $S = \{t(\mathbf{e}_1), \ldots, t(\mathbf{e}_n)\}$ that is a linear combination of the other vectors in S, then discard $\mathbf{v}$ to give the set $S_1 = S - \{\mathbf{v}\}$.
4. If there is a vector $\mathbf{v}_1$ in S_1 such that $\mathbf{v}_1$ is a linear combination of the other vectors in S_1, then discard $\mathbf{v}_1$ to give the set $S_2 = S_1 - \{\mathbf{v}_1\}$.

Continue discarding vectors in this way until you obtain a linearly independent set. This set is a basis for $\operatorname{Im} t$.

Once we know a basis for the image set of a linear transformation, we know everything that we need to know about the image set; in particular, we know its dimension.

Worked Exercise C59

Let t be the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x + 2y + 3z, 4x + y - 2z). \end{aligned}$$

Find a basis for $\operatorname{Im} t$ and state the dimension of $\operatorname{Im} t$.

Solution

We use Strategy C17.

We take the standard basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for the domain $\mathbb{R}^3$.

We determine the images of these basis vectors:

$$t(1, 0, 0) = (1, 4), \quad t(0, 1, 0) = (2, 1), \quad t(0, 0, 1) = (3, -2).$$

The set $\{(1, 4), (2, 1), (3, -2)\}$ is not linearly independent since we have

$$(3, -2) = 2(2, 1) - (1, 4),$$

so we discard $(3, -2)$ to give the set $\{(1, 4), (2, 1)\}$.

We could have discarded $(1, 4)$ or $(2, 1)$ instead: the remaining pairs of vectors are linearly independent in each case since they are not multiples of each other.

The vectors $(1, 4)$ and $(2, 1)$ are linearly independent, so $\{(1, 4), (2, 1)\}$ is a basis for $\operatorname{Im} t$.

Since the basis has two elements, $\operatorname{Im} t$ is a two-dimensional subspace of the codomain $\mathbb{R}^2$.

In fact, since $\operatorname{Im} t$ is a two-dimensional subspace of $\mathbb{R}^2$, it is the whole of $\mathbb{R}^2$.

Exercise C105

For each of the following linear transformations t, find a basis for $\operatorname{Im} t$ and state the dimension of $\operatorname{Im} t$.

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x, 2x + y)$

(b) $t : P_3 \longrightarrow P_2$
$p(x) \longmapsto p'(x)$

(c) $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$
$(x, y, z) \longmapsto (x + 2y + 3z, x + z, x + y + 2z)$

For the linear transformation t in Exercise C105(c), $\operatorname{Im} t$ is a two-dimensional subspace of the codomain $\mathbb{R}^3$. Thus $\operatorname{Im} t$ is a plane through the origin with equation

$$ax + by + cz = 0,$$

for some $a, b, c \in \mathbb{R}$ not all zero. It is possible to use the basis that you found for $\operatorname{Im} t$ in Exercise C105(c) to work out the values of a, b and c. For example, using the basis $\{(1, 1, 1), (2, 0, 1)\}$ for $\operatorname{Im} t$, we can proceed as follows. Since the basis vectors belong to $\operatorname{Im} t$, the values a, b and c satisfy the system

$$\begin{aligned} a + b + c &= 0 \\ 2a \phantom{{}+b} + c &= 0. \end{aligned}$$

The second equation gives $c = -2a$. Substituting this into the first equation gives $b = a$. So $\operatorname{Im} t$ is the plane with equation

$$ax + ay - 2az = 0$$

or, equivalently,

$$x + y - 2z = 0.$$

Finally, we note that a linear transformation $t : V \longrightarrow W$ is onto when every element of W is the image of an element of V; that is, a linear transformation is onto if and only if $\operatorname{Im} t = W$. Since $\operatorname{Im} t$ is a subspace of W, if $\dim(\operatorname{Im} t) = \dim W$ then we can immediately conclude that $\operatorname{Im} t = W$ and, conversely, if $\operatorname{Im} t = W$ then $\dim(\operatorname{Im} t) = \dim W$. Thus we have the following result.

Proposition C49

A linear transformation t is onto if and only if $\dim(\operatorname{Im} t) = \dim W$.

Exercise C106

Which of the following linear transformations are onto?

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x, 2x + y)$

(b) $t : P_3 \longrightarrow P_2$
$p(x) \longmapsto p'(x)$

(c) $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$
$(x, y, z) \longmapsto (x + 2y + 3z, x + z, x + y + 2z)$

(These are the linear transformations from Exercise C105.)

4.2 Kernel of a linear transformation

You have seen how to find the image set of a linear transformation $t : V \longrightarrow W$. Now suppose that $\mathbf{w}$ belongs to the image set of t. How can we find all the vectors in V that map to $\mathbf{w}$?

We begin by looking at the case when $\mathbf{w}$ is the zero vector. We know that $t(\mathbf{0}) = \mathbf{0}$, but it is possible that there are also some non-zero vectors in V that are mapped to $\mathbf{0}$.

For example, let t be the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x, y, 0). \end{aligned}$$

Then $t(x, y, z) = \mathbf{0}$ if and only if $(x, y, 0) = (0, 0, 0)$; that is, if and only if $x = 0$ and $y = 0$.

Thus the set of vectors that are mapped to $\mathbf{0}$ is the whole of the z-axis. This set is a one-dimensional subspace of the domain $\mathbb{R}^3$. We call this set the *kernel* of t.

The first use of *kernel* in the context of algebra was by the Russian mathematician Lev Semyonovich Pontryagin (1908–1988) in 1931. Pontryagin, who lost his eyesight in an accident when he was fourteen, was one of the leading Russian mathematicians of the twentieth century. He made fundamental contributions to algebra, topology, and dynamical systems.

Pontryagin's choice of the term *kernel* appears unrelated to its use in other areas of mathematics (integral equations, Fourier analysis).

Lev Semyonovich Pontryagin

Definition

The **kernel** of a linear transformation $t : V \longrightarrow W$ is the set

$$\operatorname{Ker} t = \{\mathbf{v} \in V : t(\mathbf{v}) = \mathbf{0}\}.$$

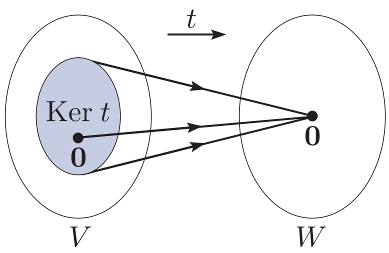


Figure 29 The kernel maps to the zero vector

Figure 29 illustrates that the kernel is the set of vectors of V mapping to the zero vector of W.

Exercise C107

Give a geometric description of the kernel of each of the following linear transformations. In each case, state whether the kernel is a subspace of the domain.

(a) $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$
$(x, y, z) \longmapsto (x, 0)$

(b) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (x, x)$

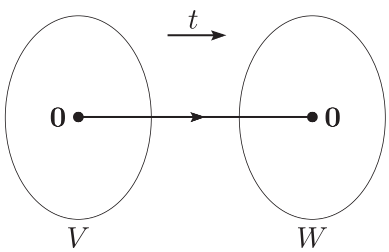


Figure 30 The vector $\mathbf{0}$ is in the kernel

For each of the linear transformations in Exercise C107, the kernel is a subspace of the domain. This is true for all linear transformations.

Theorem C50

Let $t : V \longrightarrow W$ be a linear transformation. Then $\operatorname{Ker} t$ is a subspace of the domain V.

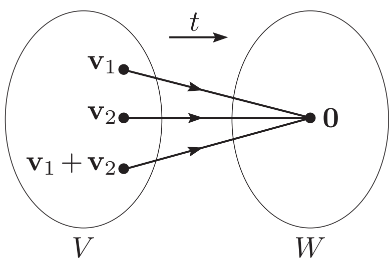


Figure 31 The kernel is closed under vector addition

Proof We use Strategy C10 in Unit C2.

First we show that $\mathbf{0} \in \operatorname{Ker} t$.

Since t is a linear transformation, $t(\mathbf{0}) = \mathbf{0}$, so $\mathbf{0} \in \operatorname{Ker} t$.

This is illustrated in Figure 30.

Next we show that $\operatorname{Ker} t$ is closed under vector addition.

Let $\mathbf{v}_1, \mathbf{v}_2 \in \operatorname{Ker} t$. Since t is a linear transformation,

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

so $\mathbf{v}_1 + \mathbf{v}_2 \in \operatorname{Ker} t$, as required.

This is illustrated in Figure 31.

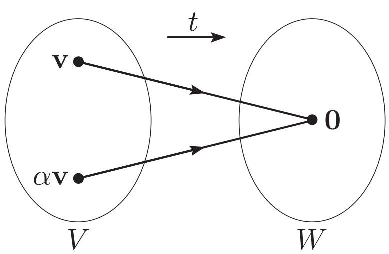


Figure 32 The kernel is closed under scalar multiplication

Finally, we show that $\operatorname{Ker} t$ is closed under scalar multiplication.

Let $\mathbf{v} \in \operatorname{Ker} t$ and $\alpha \in \mathbb{R}$. Since t is a linear transformation,

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}) = \alpha\mathbf{0} = \mathbf{0},$$

so $\alpha\mathbf{v} \in \operatorname{Ker} t$.

This is illustrated in Figure 32.

Thus $\operatorname{Ker} t$ is a subspace of V. ∎

When finding the kernel of a linear transformation, we often need to solve a system of linear equations; this sometimes involves using Gauss–Jordan elimination as in Unit C1.

Worked Exercise C60

Find the kernel and the dimension of the kernel of the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x + 2y + 3z, 4x + y - 2z). \end{aligned}$$

Solution

The kernel is the set of vectors (x, y, z) in $\mathbb{R}^3$ that satisfy

$$t(x, y, z) = \mathbf{0},$$

that is,

$$(x + 2y + 3z, 4x + y - 2z) = (0, 0).$$

Equating coordinates, we obtain the system

$$\begin{aligned} x + 2y + 3z &= 0 \\ 4x + y - 2z &= 0. \end{aligned}$$

To solve this system we row-reduce the augmented matrix.

$$\begin{array}{ll} \begin{matrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{matrix} & \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 1 & -2 & 0 \end{array}\right) \begin{matrix} 6 \\ 3 \end{matrix} \\ \mathbf{r}_2 \to \mathbf{r}_2 - 4\mathbf{r}_1 & \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -7 & -14 & 0 \end{array}\right) \begin{matrix} 6 \\ -21 \end{matrix} \\ \mathbf{r}_2 \to -\frac{1}{7}\mathbf{r}_2 & \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \begin{matrix} 6 \\ 3 \end{matrix} \\ \mathbf{r}_1 \to \mathbf{r}_1 - 2\mathbf{r}_2 & \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \begin{matrix} 0 \\ 3 \end{matrix} \end{array}$$

The augmented matrix is in row-reduced form and we have

$$\begin{aligned} x - z &= 0 \\ y + 2z &= 0. \end{aligned}$$

Assigning the parameter k to the unknown z, we obtain

$$x = k, \quad y = -2k, \quad z = k.$$

So the kernel of t is

$$\operatorname{Ker} t = \{(k, -2k, k) : k \in \mathbb{R}\};$$

that is, $\operatorname{Ker} t$ is the line through $(0, 0, 0)$ and $(1, -2, 1)$.

Thus $\operatorname{Ker} t$ is a one-dimensional subspace of the domain $\mathbb{R}^3$.

Exercise C108

For each of the following linear transformations t, find the kernel of t and the dimension of the kernel.

(a) $$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, 2x + y) \end{aligned}$$

(b) $$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x + 2y + 3z, x + z, x + y + 2z) \end{aligned}$$

We now look at examples involving vector spaces of polynomials.

Worked Exercise C61

Find the kernel of the linear transformation

$$\begin{aligned} t : P_3 &\longrightarrow P_3 \\ p(x) &\longmapsto p(x) + p(2). \end{aligned}$$

Solution

Let $p(x) = a + bx + cx^2$ be a polynomial in P_3. Then

$$\begin{aligned} t(p(x)) &= a + bx + cx^2 + a + 2b + 4c \\ &= 2a + 2b + 4c + bx + cx^2. \end{aligned}$$

The kernel of t is the set of polynomials in P_3 that satisfy $t(p(x)) = \mathbf{0}$; that is,

$$2a + 2b + 4c + bx + cx^2 = 0, \quad \text{for all } x \in \mathbb{R}.$$

Equating coefficients, we obtain the system

$$\begin{aligned} 2a + 2b + 4c &= 0 \\ b &= 0 \\ c &= 0. \end{aligned}$$

Substituting $b = 0$ and $c = 0$ into the first equation gives $a = 0$. So the only solution is $a = 0$, $b = 0$ and $c = 0$.

Thus the only polynomial in the kernel of t is the zero polynomial $p(x) = 0$; that is,

$$\operatorname{Ker} t = \{\mathbf{0}\}.$$

The kernel comprises just the zero vector so it has dimension 0.

Exercise C109

Find the kernel and dimension of the kernel of the linear transformation

$$\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x). \end{aligned}$$

For a given linear transformation $t : V \longrightarrow W$, we know how to find all the vectors in V that map to $\mathbf{0}$ in W. Now suppose that $\mathbf{b}$ ($\neq \mathbf{0}$) is some particular vector in W. How do we find all the vectors in V that map to $\mathbf{b}$? This is illustrated in Figure 33. There is a close relationship between the vectors that map to $\mathbf{b}$ and those that map to $\mathbf{0}$: if we know *one* vector $\mathbf{a}$ in V that maps to $\mathbf{b}$, that is, $t(\mathbf{a}) = \mathbf{b}$, then *every* vector $\mathbf{x}$ in V that maps to $\mathbf{b}$ may be written in the form $\mathbf{x} = \mathbf{a} + \mathbf{k}$, where $t(\mathbf{k}) = \mathbf{0}$, that is, $\mathbf{k} \in \operatorname{Ker} t$. We state this powerful result formally in the following theorem; the proof is short and constructive.

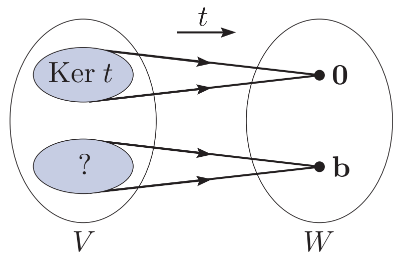


Figure 33 The vectors of V mapping to $\mathbf{b}$

Theorem C51 Solution Set Theorem

Let $t : V \longrightarrow W$ be a linear transformation. Let $\mathbf{b} \in W$ and let $\mathbf{a}$ be one vector in V that maps to $\mathbf{b}$, that is, $t(\mathbf{a}) = \mathbf{b}$. Then the solution set of the equation $t(\mathbf{x}) = \mathbf{b}$ is

$$\{\mathbf{x} : \mathbf{x} = \mathbf{a} + \mathbf{k} \text{ for some } \mathbf{k} \in \operatorname{Ker} t\}.$$

Proof The proof is in two parts. We first show that the given set is a subset of the solution set.

First we show that each vector $\mathbf{x}$ of the given form is a solution of $t(\mathbf{x}) = \mathbf{b}$. Let $\mathbf{x} = \mathbf{a} + \mathbf{k}$, where $\mathbf{k} \in \operatorname{Ker} t$. Then

$$t(\mathbf{x}) = t(\mathbf{a} + \mathbf{k}) = t(\mathbf{a}) + t(\mathbf{k}) = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

We now show that the solution set is a subset of the given set.

Conversely, we show that each vector $\mathbf{x}$ in the solution set has the given form. Let $t(\mathbf{x}) = \mathbf{b}$, where $\mathbf{x} \in V$. Then

$$t(\mathbf{x} - \mathbf{a}) = t(\mathbf{x}) - t(\mathbf{a}) = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

so $\mathbf{x} - \mathbf{a} \in \operatorname{Ker} t$; that is, $\mathbf{x} = \mathbf{a} + \mathbf{k}$, for some $\mathbf{k} \in \operatorname{Ker} t$. ■

Finally, we recall that a linear transformation $t : V \longrightarrow W$ is one-to-one if and only if no two elements in V have the same image. Thus we have the following result.

Proposition C52

A linear transformation t is one-to-one if and only if $\operatorname{Ker} t = \{\mathbf{0}\}$.

Exercise C110

Which of the following linear transformations are one-to-one?

(a) $$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, 2x + y) \end{aligned}$$

(b) $$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x + 2y + 3z, x + z, x + y + 2z) \end{aligned}$$

(c) $$\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x) \end{aligned}$$

(You found the kernels of these in Exercises C108 and C109.)

4.3 Dimension Theorem

You have seen that a linear transformation $t : V \longrightarrow W$ has two particular subspaces associated with it: $\operatorname{Ker} t$ in the domain V and $\operatorname{Im} t$ in the codomain W, as show in Figure 34.

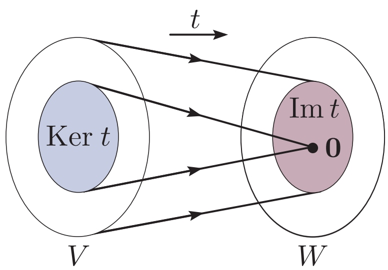


Figure 34 The subspaces $\operatorname{Ker} t$ and $\operatorname{Im} t$

There is a remarkable connection between the dimensions of these two subspaces and the dimension of the domain V.

Let t be the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x, y, 0). \end{aligned}$$

You have seen that for this linear transformation:

- the image set of t is the (x, y)-plane, so $\dim(\operatorname{Im} t) = 2$
- the kernel of t is the z-axis, so $\dim(\operatorname{Ker} t) = 1$.

Thus

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = 2 + 1 = 3,$$

which is the dimension of the domain $\mathbb{R}^3$.

Now let t be the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x + 2y + 3z, 4x + y - 2z). \end{aligned}$$

You have seen that for this linear transformation:

- the image set of t is the whole of $\mathbb{R}^2$, so $\dim(\operatorname{Im} t) = 2$
- the kernel of t is the line through $(0, 0, 0)$ and $(1, -2, 1)$, so $\dim(\operatorname{Ker} t) = 1$.

Thus

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = 2 + 1 = 3,$$

which is the dimension of the domain $\mathbb{R}^3$.

Exercise C111

For each of the following linear transformations t, calculate

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t)$$

and compare your answer with the dimension of the domain of t.

(a) $\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, 2x + y) \end{aligned}$

(b) $\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x) \end{aligned}$

(c) $\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x + 2y + 3z, x + z, x + y + 2z) \end{aligned}$

(You found the bases and dimensions of the image sets in Exercise C105, and the kernels and dimensions of the kernels in Exercises C108 and C109.)

For each of the linear transformations in Exercise C111, the dimension of the image set plus the dimension of the kernel is equal to the dimension of the domain. This relationship holds for all linear transformations. We state this result in the next theorem; if you are short of time you should skim through this proof and come back to it when time permits.

Theorem C53 Dimension Theorem

Let $t : V \longrightarrow W$ be a linear transformation. Then

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = \dim V.$$

Proof Let $\dim V = n$ and $\dim(\operatorname{Ker} t) = k$.

We show that $\dim(\operatorname{Im} t) = n - k$.

Let $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ be a basis for $\operatorname{Ker} t$. We can extend this basis, by Theorem C26 in Unit C2, to give a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for V. We prove that

$$F = \{t(\mathbf{e}_{k+1}), \dots, t(\mathbf{e}_n)\}$$

is a basis for $\operatorname{Im} t$, which shows that $\dim(\operatorname{Im} t) = n - k$.

A diagram helps here: see Figure 35.

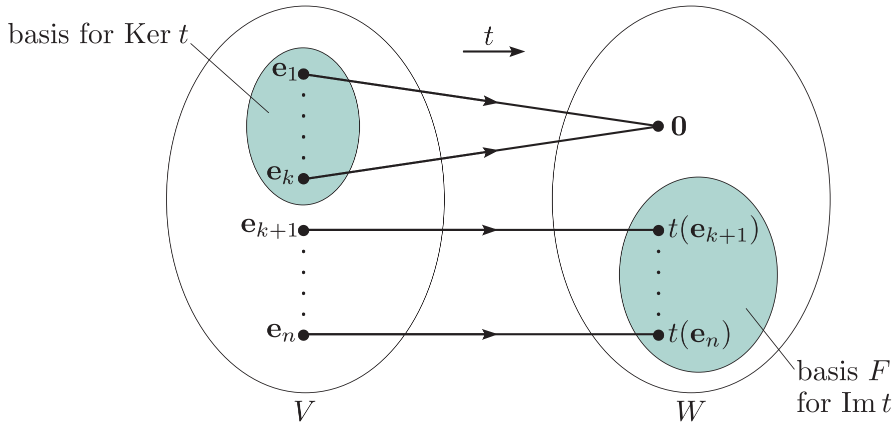


Figure 35 Bases for Ker t in V and Im t in W

To show that F is a basis for Im t, we use Strategy C8 in Unit C2.

First we prove that F spans Im t. We know from Subsection 4.1 that $\{t(\mathbf{e}_1), \ldots, t(\mathbf{e}_n)\}$ spans Im t. Since $\mathbf{e}_1, \ldots, \mathbf{e}_k$ belong to Ker t, we know that

$$t(\mathbf{e}_1) = t(\mathbf{e}_2) = \cdots = t(\mathbf{e}_k) = \mathbf{0},$$

so the span of $\{t(\mathbf{e}_1), \ldots, t(\mathbf{e}_n)\}$ is equal to the span of $\{t(\mathbf{e}_{k+1}), \ldots, t(\mathbf{e}_n)\}$. Thus F spans Im t.

Next we show that F is a linearly independent set. We must show that if

$$\alpha_{k+1}t(\mathbf{e}_{k+1}) + \alpha_{k+2}t(\mathbf{e}_{k+2}) + \cdots + \alpha_n t(\mathbf{e}_n) = \mathbf{0},$$

then

$$\alpha_{k+1} = \alpha_{k+2} = \cdots = \alpha_n = 0.$$

Since t is a linear transformation, we have

$$\alpha_{k+1}t(\mathbf{e}_{k+1}) + \cdots + \alpha_n t(\mathbf{e}_n) = t(\alpha_{k+1}\mathbf{e}_{k+1} + \cdots + \alpha_n\mathbf{e}_n).$$

So if

$$\alpha_{k+1}t(\mathbf{e}_{k+1}) + \cdots + \alpha_n t(\mathbf{e}_n) = \mathbf{0},$$

then

$$t(\alpha_{k+1}\mathbf{e}_{k+1} + \cdots + \alpha_n\mathbf{e}_n) = \mathbf{0}.$$

Thus

$$\alpha_{k+1}\mathbf{e}_{k+1} + \cdots + \alpha_n\mathbf{e}_n \in \operatorname{Ker} t.$$

Since $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ is a basis for Ker t, there exist real numbers $\alpha_1, \ldots, \alpha_k$ such that

$$\alpha_{k+1}\mathbf{e}_{k+1} + \cdots + \alpha_n\mathbf{e}_n = \alpha_1\mathbf{e}_1 + \cdots + \alpha_k\mathbf{e}_k,$$

so

$$\alpha_1\mathbf{e}_1 + \cdots + \alpha_k\mathbf{e}_k - \alpha_{k+1}\mathbf{e}_{k+1} - \cdots - \alpha_n\mathbf{e}_n = \mathbf{0}.$$

Since $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for V and so is linearly independent, it follows that

$$\alpha_1 = \cdots = \alpha_k = -\alpha_{k+1} = \cdots = -\alpha_n = 0.$$

Thus

$$\alpha_{k+1} = \alpha_{k+2} = \cdots = \alpha_n = 0,$$

as required.

Thus F is a basis for $\operatorname{Im} t$, so $\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = \dim V$. ■

The Dimension Theorem is an important result and has several applications. For example, using the Dimension Theorem we can obtain information on whether a linear transformation $t : V \longrightarrow W$ is one-to-one and/or onto.

Propositions C49 and C52 state that:

- t is onto if and only if $\dim(\operatorname{Im} t) = \dim W$
- t is one-to-one if and only if $\operatorname{Ker} t = \{\mathbf{0}\}$.

Suppose that $t : V \longrightarrow W$ is a linear transformation from the n-dimensional vector space V to the m-dimensional vector space W, as illustrated in Figure 36.

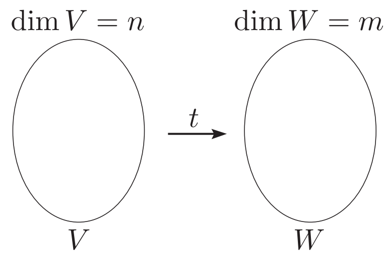


Figure 36 A linear transformation from V to W

We consider the three cases: $n > m$, $n < m$ and $n = m$.

Case (a): $n > m$

Since the image set of t is a subspace of W, we have $\dim(\operatorname{Im} t) \leq m$. It follows from the Dimension Theorem that

$$\dim(\operatorname{Ker} t) = \dim V - \dim(\operatorname{Im} t) \geq n - m > 0.$$

Thus $\operatorname{Ker} t \neq \{\mathbf{0}\}$, so t is not one-to-one, as illustrated in Figure 37.

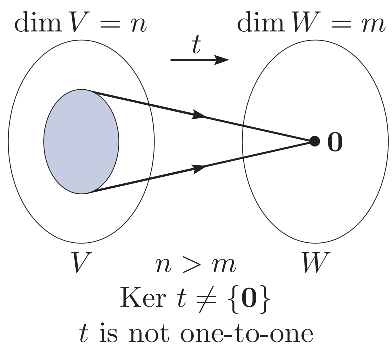


Figure 37 The case $\dim V > \dim W$

For example, the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (2x + y, x + z) \end{aligned}$$

is not one-to-one, since the dimension of the codomain (which is 2) is less than the dimension of the domain (which is 3).

This linear transformation is onto because $\dim(\operatorname{Im} t) = 2 = \dim \mathbb{R}^2$. However, in general, a linear transformation with $n > m$ may or may not be onto.

Case (b): $n < m$

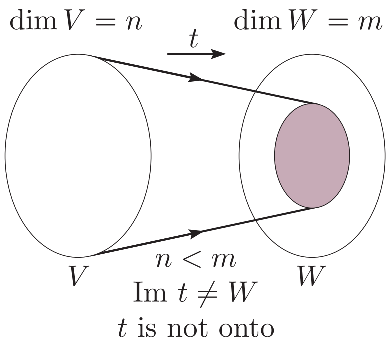


Figure 38 The case $\dim V < \dim W$

By the Dimension Theorem,

$$\dim(\operatorname{Im} t) = \dim V - \dim(\operatorname{Ker} t) \leq n < m.$$

Thus $\operatorname{Im} t$ is not the whole of the m-dimensional vector space W, so t is not onto, as illustrated in Figure 38.

For example, the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longmapsto (2x, x + y, y) \end{aligned}$$

is not onto, since the dimension of the codomain (which is 3) is greater than the dimension of the domain (which is 2).

This linear transformation is one-to-one because $\dim(\operatorname{Im} t) = 2 = \dim \mathbb{R}^2$. However, in general, a linear transformation with $n < m$ may or may not be one-to-one.

Case (c): $n = m$

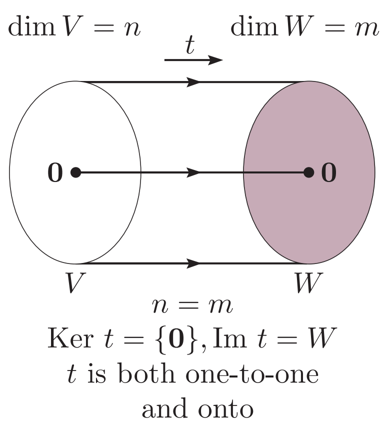


Figure 39 The case $\dim V = \dim W$ and $\operatorname{Ker} t = \{\mathbf{0}\}$

By the Dimension Theorem,

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = \dim V = n = m.$$

There are two possibilities:

- $\dim(\operatorname{Ker} t) = 0 \quad \text{and} \quad \dim(\operatorname{Im} t) = n = m$
- $\dim(\operatorname{Ker} t) > 0 \quad \text{and} \quad \dim(\operatorname{Im} t) < m.$

If $\dim(\operatorname{Ker} t) = 0$ and $\dim(\operatorname{Im} t) = n = m$, then

$$\operatorname{Ker} t = \{\mathbf{0}\} \quad \text{and} \quad \operatorname{Im} t = W.$$

Thus t is both one-to-one and onto, as illustrated in Figure 39.

For example, consider the linear transformation from Exercise C105(a),

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, 2x + y). \end{aligned}$$

Here the domain and codomain both have dimension 2, $\dim(\operatorname{Ker} t) = 0$ and $\dim(\operatorname{Im} t) = 2$. The latter is equal to the dimension of the codomain, so t is both one-to-one and onto.

If, on the other hand, $\dim(\operatorname{Ker} t) > 0$ and $\dim(\operatorname{Im} t) < m$, then

$$\operatorname{Ker} t \neq \{\mathbf{0}\} \quad \text{and} \quad \operatorname{Im} t \text{ is not the whole of } W.$$

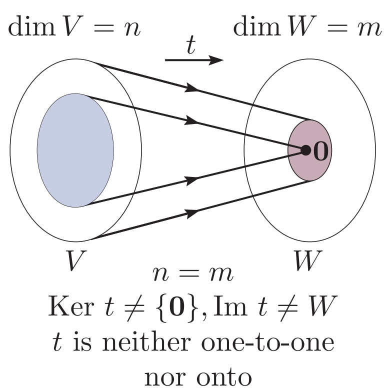


Figure 40 The case $\dim V = \dim W$ and $\operatorname{Ker} t \neq \{\mathbf{0}\}$

Thus t is neither one-to-one nor onto, as illustrated in Figure 40.

For example, consider the linear transformation from Exercise C105(c),

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x + 2y + 3z, x + z, x + y + 2z). \end{aligned}$$

Here the domain and codomain both have dimension 3, $\dim(\operatorname{Ker} t) = 1$ and $\dim(\operatorname{Im} t) = 2$. The latter is less than the dimension of the codomain of t, and thus t is neither one-to-one nor onto.

We summarise these findings in the following theorem.

Theorem C54

Let $t : V \longrightarrow W$ be a linear transformation from an n-dimensional vector space V to an m-dimensional vector space W.

(a) If $n > m$, then t is not one-to-one: $\operatorname{Ker} t \neq \{\mathbf{0}\}$.

(b) If $n < m$, then t is not onto: $\operatorname{Im} t \neq W$.

(c) If $n = m$, then

- *either* t is both one-to-one and onto:
  $$\operatorname{Ker} t = \{\mathbf{0}\} \text{ and } \operatorname{Im} t = W$$
- *or* t is neither one-to-one nor onto:
  $$\operatorname{Ker} t \neq \{\mathbf{0}\} \text{ and } \operatorname{Im} t \neq W.$$

Exercise C112

What can we deduce from Theorem C54 about the following linear transformations?

(a) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$
$(x, y) \longmapsto (x, y, x + y)$

(b) $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$
$(x, y) \longmapsto (3x, 4x + y)$

(c) $t : P_3 \longrightarrow P_2$
$p(x) \longmapsto p'(x)$

Systems of linear equations

You will now see how we can use linear transformations to obtain information on the number of solutions of a system of linear equations.

Suppose that we want to know how many solutions there are to the following system of three linear equations in three unknowns:

$$\begin{aligned} 2x + 3y + 4z &= 7 \\ x + 5y + 6z &= 4 \\ 3x + 2y + 5z &= 1. \end{aligned}$$

This system can be written in matrix form as

$$\begin{pmatrix} 2 & 3 & 4 \\ 1 & 5 & 6 \\ 3 & 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ 1 \end{pmatrix}.$$

Now let t be the linear transformation with the matrix representation

$$t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 3 & 4 \\ 1 & 5 & 6 \\ 3 & 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

We see that (x, y, z) is a solution of the system of equations precisely when $t(x, y, z) = (7, 4, 1)$.

Thus the number of solutions of the system of equations is the same as the number of vectors in $\mathbb{R}^3$ that map to the vector $(7, 4, 1)$ under t.

In general, suppose that we want to know how many solutions there are to the system of m linear equations in n unknowns with the matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

Let t be the linear transformation with the matrix representation

$$\begin{aligned} t : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \mathbf{x} &\longmapsto \mathbf{A}\mathbf{x}. \end{aligned}$$

Then the number of solutions of the system of equations is the same as the number of vectors that map to $\mathbf{b}$ under t.

Suppose $\mathbf{b} \in \operatorname{Im} t$. Then there is some vector $\mathbf{a} \in \mathbb{R}^n$ such that $t(\mathbf{a}) = \mathbf{b}$. Then, using the Solution Set Theorem (Theorem C51), the solution set to the system of equations is

$$\{\mathbf{x} : \mathbf{x} = \mathbf{a} + \mathbf{k} \text{ for some } \mathbf{k} \in \operatorname{Ker} t\}.$$

Now $\operatorname{Ker} t$ is a subspace of $\mathbb{R}^n$, by Theorem C50. A subspace of $\mathbb{R}^n$ of dimension 0 comprises just the zero vector. A subspace of $\mathbb{R}^n$ of dimension greater than 0 comprises infinitely many vectors since it is a line, a plane or a higher-dimensional space. So $\operatorname{Ker} t$ contains either just the zero vector or infinitely many vectors.

It follows that there are three possibilities for the number of solutions:

- if $\mathbf{b} \in \operatorname{Im} t$ and $\operatorname{Ker} t = \{\mathbf{0}\}$, then there is exactly *one* solution
- if $\mathbf{b} \in \operatorname{Im} t$ and $\operatorname{Ker} t \neq \{\mathbf{0}\}$, then there are *infinitely many* solutions
- if $\mathbf{b} \notin \operatorname{Im} t$, then there are *no* solutions.

Thus a system of linear equations has no solutions, or one solution, or infinitely many solutions. This result was stated without proof in Subsection 1.2 of Unit C1.

Exercise C113

How many solutions are there to the following system of three linear equations in three unknowns?

$$\begin{aligned} x + 2y + 3z &= 1 \\ x \phantom{{}+ 2y} + z &= 1 \\ x + y + 2z &= 1 \end{aligned}$$

Use your solutions to Exercises C105(c) and C108(b).

By considering the linear transformation

$$\begin{aligned} t : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \mathbf{x} &\longmapsto \mathbf{A}\mathbf{x}, \end{aligned}$$

we can show that the number of solutions of the system $\mathbf{Ax} = \mathbf{b}$ of m linear equations in n unknowns depends on the values of m and n. We consider three cases: $n > m$, $n < m$ and $n = m$.

Case (a): $n > m$

It follows from Theorem C54 that $\operatorname{Ker} t \neq \{\mathbf{0}\}$. Thus the equation $\mathbf{Ax} = \mathbf{b}$ has either no solution (if $\mathbf{b} \notin \operatorname{Im} t$) or infinitely many solutions (if $\mathbf{b} \in \operatorname{Im} t$). For example, the system

$$\begin{aligned} 2x + y + z &= a \\ 4x + 2y + 2z &= b, \end{aligned}$$

of two equations in three unknowns has either no solution or infinitely many solutions, depending on the values of a and b. For example, the system has no solution when $a = 3$ and $b = 4$, and infinitely many solutions when $a = 2$ and $b = 4$.

Case (b): $n < m$

It follows from Theorem C54 that $\operatorname{Im} t \neq \mathbb{R}^m$. Thus there is some $\mathbf{b}$ for which the equation $\mathbf{Ax} = \mathbf{b}$ has no solutions. For example, there are some values of a, b and c for which the system

$$\begin{aligned} 2x + y &= a \\ x + 3y &= b \\ 4x + y &= c, \end{aligned}$$

of three equations in two unknowns has no solutions. For example, the system has no solutions when $a = 3$, $b = 4$ and $c = 2$.

Case (c): $n = m$

It follows from Theorem C54 that there are two possibilities.

If $\operatorname{Ker} t = \{\mathbf{0}\}$ and $\operatorname{Im} t = \mathbb{R}^m$, then the equation $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for each $\mathbf{b}$. For example, the system

$$\begin{aligned} x + y &= a \\ y &= b, \end{aligned}$$

of two equations in two unknowns has exactly one solution, namely $(x, y) = (a - b, b)$, for each pair of values (a, b).

If $\operatorname{Ker} t \neq \{\mathbf{0}\}$ and $\operatorname{Im} t \neq \mathbb{R}^m$, then there exist vectors $\mathbf{b}$ for which the equation $\mathbf{Ax} = \mathbf{b}$ has no solutions, and for all other $\mathbf{b}$, the equation $\mathbf{Ax} = \mathbf{b}$ has infinitely many solutions. Consider the system

$$\begin{aligned} x + 2y &= a \\ 2x + 4y &= b, \end{aligned}$$

of two equations in two unknowns. Since $2x + 4y = 2(x + 2y)$, these equations have no solution when $b \neq 2a$. When $b = 2a$, however, putting $y = k$ gives $(x, y) = (a - 2k, k)$, where $k \in \mathbb{R}$, as a solution of the equations; thus there are infinitely many solutions.

We summarise these results below.

Theorem C55

Let $\mathbf{Ax} = \mathbf{b}$ be a system of m linear equations in n unknowns.

(a) If $n > m$, then $\mathbf{Ax} = \mathbf{b}$ has either no solution or infinitely many solutions.

(b) If $n < m$, then there is some $\mathbf{b}$ for which $\mathbf{Ax} = \mathbf{b}$ has no solutions.

(c) If $n = m$, then:

- *either* $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for each $\mathbf{b}$
- *or* there are some $\mathbf{b}$ for which $\mathbf{Ax} = \mathbf{b}$ has no solutions; for all other $\mathbf{b}$, $\mathbf{Ax} = \mathbf{b}$ has infinitely many solutions.

Exercise C114

What can you deduce from Theorem C55 about the number of solutions of each of the following systems of linear equations?

(a) $$\begin{aligned} 3x + y + z &= 1 \\ 4x + 2y + 4z &= 3 \end{aligned}$$

(b) $$\begin{aligned} 3x + y + z &= a \\ 4x + 2y + 4z &= b \\ 5x + y + 6z &= c \end{aligned}$$

Summary

In this unit you have seen that linear transformations are functions between vector spaces that preserve linear combinations of vectors, and that for finite-dimensional vector spaces they are precisely the functions that have matrix representations. Using properties of matrices you have investigated invertible linear transformations. You have seen that finite-dimensional vector spaces are isomorphic if and only if their dimensions are equal, and hence that all vector spaces of dimension n are isomorphic to $\mathbb{R}^n$. You have met the Dimension Theorem, the important result that the sum of dimensions of the image set and kernel are equal to the dimension of the domain. In addition, you have seen that linear transformations can be used to prove that matrix multiplication is associative and to help determine the number of solutions of a system of linear equations.

Learning outcomes

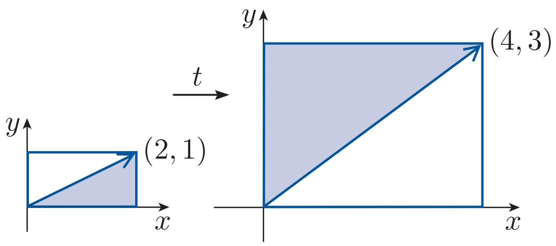
After working through this unit, you should be able to:

- explain what is meant by a *linear transformation* and understand that linear transformations preserve the zero vector and linear combinations of vectors
- recognise simple linear transformations of the plane
- determine whether or not a given function is a linear transformation
- understand that the *matrix representation* of a linear transformation $t : V \longrightarrow W$ depends on the bases used for V and W
- find the matrix representation, with respect to given bases, of a linear transformation between finite-dimensional vector spaces
- understand the relationship between matrices and linear transformations
- use the matrix representations of two given linear transformations s and t to find a matrix representation of the composite function $s \circ t$
- determine whether a given linear transformation is invertible and, if it is, find its inverse
- understand that each n-dimensional vector space is isomorphic to $\mathbb{R}^n$
- explain the meaning of the terms *image set* and *kernel* of a linear transformation
- find a basis for the image set of a given linear transformation and find the kernel of a given linear transformation
- understand the relationship between the dimension of the image set, the dimension of the kernel and the dimension of the domain of a linear transformation
- understand that the number of solutions of a system of m linear equations in n unknowns depends on the values of m and n.

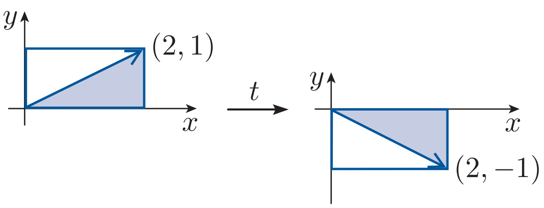
Solutions to exercises

Solution to Exercise C82

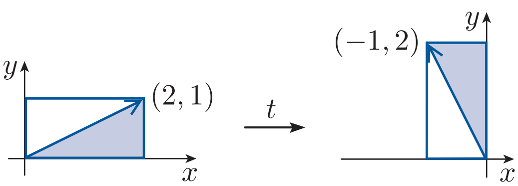
(a) This is a $(2, 3)$-scaling.



(b) This is q_0, a reflection in the x-axis; it is also a $(1, -1)$-scaling.



(c) This is $r_{\pi/2}$, a rotation through an angle $\pi/2$.



Solution to Exercise C83

(a) First $t(\mathbf{0}) = \mathbf{0}$, so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2.$$

In $\mathbb{R}^2$, let $\mathbf{v}_1 = (x_1, y_1)$ and $\mathbf{v}_2 = (x_2, y_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2) \\ &= (x_1 + x_2 + 3(y_1 + y_2), y_1 + y_2) \\ &= (x_1 + x_2 + 3y_1 + 3y_2, y_1 + y_2) \end{aligned}$$

and

$$\begin{aligned} t(\mathbf{v}_1) + t(\mathbf{v}_2) &= (x_1 + 3y_1, y_1) + (x_2 + 3y_2, y_2) \\ &= (x_1 + x_2 + 3y_1 + 3y_2, y_1 + y_2). \end{aligned}$$

These expressions are equal, so LT1 is satisfied.

Finally, we check whether t satisfies LT2:

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in \mathbb{R}^2,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} = (x, y)$ be a vector in $\mathbb{R}^2$ and let $\alpha \in \mathbb{R}$. Then

$$t(\alpha\mathbf{v}) = t(\alpha x, \alpha y) = (\alpha x + 3\alpha y, \alpha y)$$

and

$$\alpha\, t(\mathbf{v}) = \alpha(x + 3y, y) = (\alpha x + 3\alpha y, \alpha y).$$

These expressions are equal, so LT2 is satisfied.

Since both LT1 and LT2 are satisfied, t is a linear transformation.

(b) Since $t(\mathbf{0}) = t(0, 0) = (2, 1) \neq \mathbf{0}$, it follows from Strategy C14 that t is not a linear transformation.

Solution to Exercise C84

We use Strategy C14.

(a) First $t(\mathbf{0}) = \mathbf{0}$, so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2.$$

In $\mathbb{R}^2$, let $\mathbf{v}_1 = (x_1, y_1)$ and $\mathbf{v}_2 = (x_2, y_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2) \\ &= (x_1 + x_2, y_1 + y_2, x_1 + x_2, y_1 + y_2) \end{aligned}$$

and

$$\begin{aligned} t(\mathbf{v}_1) + t(\mathbf{v}_2) &= (x_1, y_1, x_1, y_1) + (x_2, y_2, x_2, y_2) \\ &= (x_1 + x_2, y_1 + y_2, x_1 + x_2, y_1 + y_2). \end{aligned}$$

These expressions are equal, so LT1 is satisfied.

Finally, we check whether t satisfies LT2:

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in \mathbb{R}^2,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} = (x, y)$ be a vector in $\mathbb{R}^2$ and let $\alpha \in \mathbb{R}$. Then

$$t(\alpha\mathbf{v}) = t(\alpha x, \alpha y) = (\alpha x, \alpha y, \alpha x, \alpha y)$$

and

$$\alpha\, t(\mathbf{v}) = \alpha(x, y, x, y) = (\alpha x, \alpha y, \alpha x, \alpha y).$$

These expressions are equal, so LT2 is satisfied.

Since both LT1 and LT2 are satisfied, t is a linear transformation.

(b) First $t(\mathbf{0}) = \mathbf{0}$, so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3.$$

In $\mathbb{R}^3$, let $\mathbf{v}_1 = (x_1, y_1, z_1)$ and $\mathbf{v}_2 = (x_2, y_2, z_2)$. Then

$$\begin{aligned} t(\mathbf{v}_1 + \mathbf{v}_2) &= t(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_1 + x_2)^2 \\ &= x_1^2 + x_2^2 + 2x_1x_2 \end{aligned}$$

and

$$t(\mathbf{v}_1) + t(\mathbf{v}_2) = x_1^2 + x_2^2.$$

These expressions are not equal in general, so LT1 is not satisfied.

Thus t is not a linear transformation.

(c) Since $t(\mathbf{0}) = t(0,0,0) = (0,0,0,1) \neq \mathbf{0}$, it follows that t is not a linear transformation.

Solution to Exercise C85

First we show that t satisfies LT1:

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3.$$

In $\mathbb{R}^3$, let $\mathbf{v}_1 = (x_1, y_1, z_1)$ and $\mathbf{v}_2 = (x_2, y_2, z_2)$. Then

$$\begin{aligned} &t(\mathbf{v}_1 + \mathbf{v}_2) \\ &= t(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= ((x_1 + x_2)\cos\theta - (y_1 + y_2)\sin\theta, \\ &\qquad (x_1 + x_2)\sin\theta + (y_1 + y_2)\cos\theta, z_1 + z_2) \end{aligned}$$

and

$$\begin{aligned} &t(\mathbf{v}_1) + t(\mathbf{v}_2) \\ &= (x_1\cos\theta - y_1\sin\theta, x_1\sin\theta + y_1\cos\theta, z_1) \\ &\qquad + (x_2\cos\theta - y_2\sin\theta, x_2\sin\theta + y_2\cos\theta, z_2) \\ &= ((x_1 + x_2)\cos\theta - (y_1 + y_2)\sin\theta, \\ &\qquad (x_1 + x_2)\sin\theta + (y_1 + y_2)\cos\theta, z_1 + z_2). \end{aligned}$$

These expressions are equal, so LT1 is satisfied.

Next we show that t satisfies LT2:

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}), \quad \text{for all } \mathbf{v} \in \mathbb{R}^3,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} = (x, y, z)$ be a vector in $\mathbb{R}^3$ and let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} t(\alpha\mathbf{v}) &= t(\alpha x, \alpha y, \alpha z) \\ &= (\alpha x\cos\theta - \alpha y\sin\theta, \alpha x\sin\theta + \alpha y\cos\theta, \alpha z) \end{aligned}$$

and

$$\begin{aligned} \alpha\, t(\mathbf{v}) &= \alpha(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z) \\ &= (\alpha x\cos\theta - \alpha y\sin\theta, \alpha x\sin\theta + \alpha y\cos\theta, \alpha z). \end{aligned}$$

These expressions are equal, so LT2 is satisfied.

Since LT1 and LT2 are satisfied, t is a linear transformation.

Solution to Exercise C86

We use Strategy C14.

Since the zero element of P_3 is $p(x) = 0$, we have $p(2) = 0$ and thus $t(\mathbf{0}) = \mathbf{0}$; so t may be a linear transformation.

Next we check whether t satisfies LT1:

$$t(p(x) + q(x)) = t(p(x)) + t(q(x)), \quad \text{for all } p(x), q(x) \in P_3.$$

Let $p(x), q(x) \in P_3$. Then

$$t(p(x) + q(x)) = p(x) + q(x) + p(2) + q(2)$$

and

$$\begin{aligned} t(p(x)) + t(q(x)) &= p(x) + p(2) + q(x) + q(2) \\ &= p(x) + q(x) + p(2) + q(2). \end{aligned}$$

These expressions are equal, so LT1 is satisfied.

Finally, we check whether t satisfies LT2:

$$t(\alpha p(x)) = \alpha\, t(p(x)), \quad \text{for all } p(x) \in P_3,\ \alpha \in \mathbb{R}.$$

Let $p(x) \in P_3$ and $\alpha \in \mathbb{R}$. Then

$$t(\alpha p(x)) = \alpha p(x) + \alpha p(2)$$

and

$$\alpha\, t(p(x)) = \alpha(p(x) + p(2)) = \alpha p(x) + \alpha p(2).$$

These expressions are equal, so LT2 is satisfied.

Since both LT1 and LT2 are satisfied, t is a linear transformation.

Solution to Exercise C87

First we show that i_V satisfies LT1:

$$i_V(\mathbf{v}_1 + \mathbf{v}_2) = i_V(\mathbf{v}_1) + i_V(\mathbf{v}_2), \quad \text{for all } \mathbf{v}_1, \mathbf{v}_2 \in V.$$

Let $\mathbf{v}_1, \mathbf{v}_2 \in V$. Then

$$i_V(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v}_1 + \mathbf{v}_2$$

and

$$i_V(\mathbf{v}_1) + i_V(\mathbf{v}_2) = \mathbf{v}_1 + \mathbf{v}_2.$$

These expressions are equal, so LT1 is satisfied.

Next we show that i_V satisfies LT2:

$$i_V(\alpha\mathbf{v}) = \alpha\, i_V(\mathbf{v}), \quad \text{for all } \mathbf{v} \in V,\ \alpha \in \mathbb{R}.$$

Let $\mathbf{v} \in V$ and $\alpha \in \mathbb{R}$. Then

$$i_V(\alpha\mathbf{v}) = \alpha\mathbf{v}$$

and

$$\alpha\, i_V(\mathbf{v}) = \alpha\mathbf{v}.$$

These expressions are equal, so LT2 is satisfied.

Since LT1 and LT2 are satisfied, t is a linear transformation.

Solution to Exercise C88

We can write any vector (x, y) in $\mathbb{R}^2$ in the form

$$(x, y) = x(1, 0) + y(0, 1).$$

It follows from Theorem C39 that

$$\begin{aligned} q_\phi(x, y) &= q_\phi(x(1, 0) + y(0, 1)) \\ &= xq_\phi(1, 0) + yq_\phi(0, 1) \\ &= x(\cos 2\phi, \sin 2\phi) + y(\sin 2\phi, -\cos 2\phi) \\ &= (x\cos 2\phi + y\sin 2\phi, x\sin 2\phi - y\cos 2\phi). \end{aligned}$$

Solution to Exercise C89

(a) Here $E = \{(3, 1), (2, 1)\}$. Therefore,

$$\mathbf{v} = (3, 1) = 1(3, 1) + 0(2, 1),$$

so

$$\mathbf{v}_E = (1, 0)_E.$$

(b) Here $E = \{(1, 2), (2, 1)\}$. We must find $a, b \in \mathbb{R}$ such that

$$(3, 1) = (a, b)_E.$$

Since

$$(a, b)_E = a(1, 2) + b(2, 1) = (a + 2b, 2a + b),$$

equating corresponding coordinates gives the system

$$\begin{aligned} a + 2b &= 3 \\ 2a + b &= 1. \end{aligned}$$

Solving, we have $a = -\frac{1}{3}$ and $b = \frac{5}{3}$, so

$$\mathbf{v}_E = \left(-\tfrac{1}{3}, \tfrac{5}{3}\right)_E.$$

Solution to Exercise C90

(a) Here $E = \{1, x\}$. Therefore,

$$p(x) = 2 + 3x = 2 \times (1) + 3 \times (x),$$

so the E-coordinate representation of $p(x)$ is

$$(2, 3)_E.$$

(b) Here $E = \{1, 4 + 6x\}$. Therefore,

$$p(x) = 2 + 3x = 0 \times (1) + \tfrac{1}{2} \times (4 + 6x),$$

so the E-coordinate representation of $p(x)$ is

$$\left(0, \tfrac{1}{2}\right)_E.$$

(c) Here $E = \{2x, 1 + 4x\}$. We must find $a, b \in \mathbb{R}$ such that

$$p(x) = 2 + 3x = (a, b)_E.$$

Since

$$\begin{aligned} (a, b)_E &= a \times 2x + b \times (1 + 4x) \\ &= b + (2a + 4b)x, \end{aligned}$$

equating corresponding coefficients gives the system

$$\begin{aligned} b &= 2 \\ 2a + 4b &= 3. \end{aligned}$$

Solving, we have $a = -\frac{5}{2}$ and $b = 2$. Thus the E-coordinate representation of $p(x)$ is

$$\left(-\tfrac{5}{2}, 2\right)_E.$$

Solution to Exercise C91

(a) We have

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \longmapsto \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

so $t(1, 0) = (3, 0)$.

Similarly,

$$\begin{pmatrix}0\\1\end{pmatrix} \longmapsto \begin{pmatrix}3&0\\0&2\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\2\end{pmatrix},$$

so $t(0,1) = (0,2)$.

Thus the coordinates of $t(1,0)$ form the first column of the matrix of t, and the coordinates of $t(0,1)$ form the second column of the matrix of t.

(b) We have

$$\begin{pmatrix}1\\0\end{pmatrix} \longmapsto \begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix},$$

so $t(1,0) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}(1,1)$.

Similarly,

$$\begin{pmatrix}0\\1\end{pmatrix} \longmapsto \begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix},$$

so $t(0,1) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}(-1,1)$.

As in part (a), the coordinates of $t(1,0)$ form the first column of the matrix of t, and the coordinates of $t(0,1)$ form the second column of the matrix of t.

Solution to Exercise C92

We use Strategy C15.

(a) We find the images of the vectors in the domain basis $E = \{(1,0),(0,1)\}$:

$$t(1,0) = (1,0), \quad t(0,1) = (3,1).$$

We find the F-coordinates of each of these image vectors, where $F = \{(1,0),(0,1)\}$:

$$t(1,0) = (1,0)_F, \quad t(0,1) = (3,1)_F.$$

Hence the matrix of t with respect to the standard bases for the domain and codomain is

$$\mathbf{A} = \begin{pmatrix}1&3\\0&1\end{pmatrix}.$$

Thus the matrix representation of t with respect to these bases is

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto \begin{pmatrix}1&3\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+3y\\y\end{pmatrix}.$$

(b) We find the images of the vectors in the domain basis $E = \{1, x, x^2\}$. The first basis vector is the constant polynomial $p_0(x) = 1$, for which $p_0(2) = 1$. The second basis vector is $p_1(x) = x$, for which $p_1(2) = 2$; and the third basis vector is $p_2(x) = x^2$, for which $p_2(2) = 4$. Thus

$$t(1) = 1 + 1 = 2, \quad t(x) = x + 2,$$
$$t(x^2) = x^2 + 2^2 = x^2 + 4.$$

We find the F-coordinates of each of these image vectors, where $F = \{1, x, x^2\}$:

$$t(1) = (2,0,0)_F, \quad t(x) = (2,1,0)_F,$$
$$t(x^2) = (4,0,1)_F.$$

Hence the matrix of t with respect to the standard bases for the domain and codomain is

$$\mathbf{A} = \begin{pmatrix}2&2&4\\0&1&0\\0&0&1\end{pmatrix}.$$

Thus the matrix representation of t with respect to these bases is

$$\begin{pmatrix}a\\b\\c\end{pmatrix} \longmapsto \begin{pmatrix}2&2&4\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix} = \begin{pmatrix}2a+2b+4c\\b\\c\end{pmatrix}.$$

(Notice that

$$\begin{aligned} t(a+bx+cx^2) &= (a+bx+cx^2) + (a+2b+2^2c)\\ &= a+bx+cx^2+a+2b+4c\\ &= 2a+2b+4c+bx+cx^2.) \end{aligned}$$

(c) We find the images of the vectors in the domain basis $E = \{(1,0),(0,1)\}$:

$$t(1,0) = (1,0,1,0), \quad t(0,1) = (0,1,0,1).$$

We find the F-coordinates of each of these image vectors, where
$F = \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$:

$$t(1,0) = (1,0,1,0)_F, \quad t(0,1) = (0,1,0,1)_F.$$

Hence the matrix of t with respect to the standard bases for the domain and codomain is

$$\mathbf{A} = \begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}.$$

Thus the matrix representation of t with respect to these bases is

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto \begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\\y\\x\\y\end{pmatrix}.$$

(d) We find the images of the vectors in the domain basis $E = \{(1,0,0),(0,1,0),(0,0,1)\}$:

$$t(1,0,0) = (1,0), \quad t(0,1,0) = (0,1),$$
$$t(0,0,1) = (0,0).$$

We find the F-coordinates of each of these image vectors, where $F = \{(1,0),(0,1)\}$:

$$t(1,0,0) = (1,0)_F, \quad t(0,1,0) = (0,1)_F$$
$$t(0,0,1) = (0,0)_F.$$

Hence the matrix of t with respect to the standard bases for the domain and codomain is

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Thus the matrix representation of t with respect to these bases is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Solution to Exercise C93

We use Strategy C15.

(a) We find the images of the vectors in the domain basis $E = \{(1,0,1),(1,0,0),(1,1,1)\}$:

$$t(1,0,1) = (1,0), \quad t(1,0,0) = (1,0),$$
$$t(1,1,1) = (1,1).$$

We find the F-coordinates of each of these image vectors, where $F = \{(1,0),(0,1)\}$:

$$t(1,0,1) = (1,0)_F, \quad t(1,0,0) = (1,0)_F,$$
$$t(1,1,1) = (1,1)_F.$$

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) We find the images of the vectors in the domain basis $E = \{(1,0,0),(0,1,0),(0,0,1)\}$:

$$t(1,0,0) = (1,0), \quad t(0,1,0) = (0,1),$$
$$t(0,0,1) = (0,0).$$

We find the F-coordinates of each of these image vectors, where $F = \{(2,1),(1,1)\}$.

For the first image vector we need $a, b \in \mathbb{R}$ such that

$$(1,0) = (a,b)_F.$$

Since

$$(a,b)_F = a(2,1) + b(1,1) = (2a+b, a+b),$$

by equating coordinates we see that $a = 1$ and $b = -1$, so $(1,0) = (1,-1)_F$. Therefore

$$t(1,0,0) = (1,-1)_F.$$

For the second image vector we need $c, d \in \mathbb{R}$ such that

$$(0,1) = (c,d)_F.$$

Since

$$(c,d)_F = c(2,1) + d(1,1) = (2c+d, c+d),$$

by equating coordinates we obtain the system

$$\begin{aligned} 2c + d &= 0 \\ c + d &= 1. \end{aligned}$$

Solving, we have $c = -1$ and $d = 2$, so $(0,1) = (-1,2)_F$. Therefore

$$t(0,1,0) = (-1,2)_F.$$

Finally, for the third image vector we need $e, f \in \mathbb{R}$ such that

$$(0,0) = (e,f)_F.$$

Using the same method as before we have $e = f = 0$, so $(0,0) = (0,0)_F$. Therefore

$$t(0,0,1) = (0,0)_F.$$

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 0 \end{pmatrix}.$$

(c) We find the images of the vectors in the domain basis $E = \{(0,1,0),(1,1,1),(0,1,1)\}$:

$$t(0,1,0) = (0,1), \quad t(1,1,1) = (1,1),$$
$$t(0,1,1) = (0,1).$$

We find the F-coordinates of each of these image vectors, where $F = \{(1,3),(2,4)\}$.

For the first image vector we need $a, b \in \mathbb{R}$ such that

$$(0,1) = (a,b)_F.$$

Since

$$(a,b)_F = a(1,3) + b(2,4) = (a+2b, 3a+4b),$$

by equating coordinates we obtain the system

$$\begin{aligned} a + 2b &= 0\\ 3a + 4b &= 1. \end{aligned}$$

Solving, we have $a = 1$ and $b = -\frac{1}{2}$, so

$$(0,1) = \left(1, -\tfrac{1}{2}\right)_F.$$

Therefore

$$t(0,1,0) = \left(1, -\tfrac{1}{2}\right)_F.$$

For the second image vector we need $c, d \in \mathbb{R}$ such that

$$(1,1) = (c,d)_F.$$

Since

$$(c,d)_F = c(1,3) + d(2,4) = (c + 2d, 3c + 4d),$$

by equating coordinates we obtain the system

$$\begin{aligned} c + 2d &= 1\\ 3c + 4d &= 1. \end{aligned}$$

Solving, we have $c = -1$ and $d = 1$, so $(1,1) = (-1,1)_F$. Therefore

$$t(1,1,1) = (-1,1)_F.$$

Since $t(0,1,1) = (0,1) = t(0,1,0)$, we have

$$t(0,1,1) = \left(1, -\tfrac{1}{2}\right)_F.$$

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 1\\ -\frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix}.$$

Solution to Exercise C94

We use Strategy C15.

(a) We find the images of the polynomials in the domain basis $E = \{1, x, x^2\}$:

$$t(1) = 0, \quad t(x) = 1, \quad t(x^2) = 2x.$$

We find the F-coordinates of each of these image vectors, where $F = \{2x, 1 + x\}$.

For the first image vector we have

$$t(1) = 0 = (0,0)_F.$$

For the second image vector we need $a, b \in \mathbb{R}$ such that

$$1 = (a,b)_F.$$

Since

$$\begin{aligned} (a,b)_F &= a \times (2x) + b \times (1 + x)\\ &= b + (2a + b)x, \end{aligned}$$

by equating coefficients we obtain the system

$$\begin{aligned} b &= 1\\ 2a + b &= 0. \end{aligned}$$

Solving, we have $a = -\frac{1}{2}$ and $b = 1$, so $1 = \left(-\frac{1}{2}, 1\right)_F$. Therefore

$$t(x) = \left(-\tfrac{1}{2}, 1\right)_F.$$

For the final image vector we have

$$t(x^2) = 2x = (1,0)_F.$$

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} 0 & -\frac{1}{2} & 1\\ 0 & 1 & 0 \end{pmatrix}.$$

Thus the matrix representation of t with respect to the standard basis E and non-standard basis F is

$$\begin{pmatrix} a\\ b\\ c \end{pmatrix}_E \longmapsto \begin{pmatrix} 0 & -\frac{1}{2} & 1\\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a\\ b\\ c \end{pmatrix}_E = \begin{pmatrix} -\frac{1}{2}b + c\\ b \end{pmatrix}_F.$$

(b) We find the images of the polynomials in the domain basis $E = \{x, x^2, 1\}$:

$$t(x) = 1, \quad t(x^2) = 2x, \quad t(1) = 0.$$

We find the F-coordinates of each of these image vectors, where $F = \{2x, 1 + x\}$.

We know from part (a) that

$$\begin{aligned} t(x) &= \left(-\tfrac{1}{2}, 1\right)_F,\\ t(x^2) &= (1,0)_F,\\ t(1) &= (0,0)_F. \end{aligned}$$

Hence the matrix of t with respect to the bases E and F is

$$\mathbf{A} = \begin{pmatrix} -\frac{1}{2} & 1 & 0\\ 1 & 0 & 0 \end{pmatrix}.$$

Thus the matrix representation of t with respect to the non-standard bases E and F is

$$\begin{pmatrix} a\\ b\\ c \end{pmatrix}_E \longmapsto \begin{pmatrix} -\frac{1}{2} & 1 & 0\\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a\\ b\\ c \end{pmatrix}_E = \begin{pmatrix} -\frac{1}{2}a + b\\ a \end{pmatrix}_F.$$

Solution to Exercise C95

The functions in parts (a) and (d) are linear transformations since they are of the form

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (ax + by, cx + dy), \end{aligned}$$

for some $a, b, c, d \in \mathbb{R}$.

The functions in parts (b) and (c) are not linear transformations since they are not of this form.

Solution to Exercise C96

(a) We have

$$\begin{aligned} r(p(x, y)) &= r(3x + y, -x) \\ &= (3x + y, (3x + y) - x) \\ &= (3x + y, 2x + y). \end{aligned}$$

Thus $r \circ p$ is given by

$$\begin{aligned} r \circ p : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (3x + y, 2x + y). \end{aligned}$$

(b) We have

$$\begin{aligned} p(r(x, y)) &= p(x, x + y) \\ &= (3x + (x + y), -x) \\ &= (4x + y, -x). \end{aligned}$$

Thus $p \circ r$ is given by

$$\begin{aligned} p \circ r : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (4x + y, -x). \end{aligned}$$

Solution to Exercise C97

It follows from the Composition Rule that the matrix of $s \circ t$ with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 4 \\ 2 & 1 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 4 & 11 \\ 4 & 2 & 0 & 6 \\ 1 & 0 & 2 & 4 \end{pmatrix}.$$

Thus the matrix representation of $s \circ t$ with respect to the standard bases for the domain and codomain is

$$s \circ t : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$$

$$\begin{aligned} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} &\longmapsto \begin{pmatrix} 4 & 1 & 4 & 11 \\ 4 & 2 & 0 & 6 \\ 1 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \\ &= \begin{pmatrix} 4x + y + 4z + 11w \\ 4x + 2y + 6w \\ x + 2z + 4w \end{pmatrix}. \end{aligned}$$

Solution to Exercise C98

It follows from the Composition Rule that the matrix of $s \circ t$ with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 2 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Thus the matrix representation of $s \circ t$ with respect to the standard bases for the domain and codomain is

$$s \circ t : P_3 \longrightarrow P_2$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}_E \longmapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}_E = \begin{pmatrix} b \\ 2c \end{pmatrix}_F.$$

As expected, this is the same as the matrix representation for s.

Solution to Exercise C99

Since

$$\begin{aligned} &s(t(x, y)) \\ &\quad = s(4x - y, -3x + y) \\ &\quad = ((4x - y) + (-3x + y), 3(4x - y) + 4(-3x + y)) \\ &\quad = (x, y) \end{aligned}$$

and

$$\begin{aligned} &t(s(x, y)) \\ &\quad = t(x + y, 3x + 4y) \\ &\quad = (4(x + y) - (3x + 4y), -3(x + y) + (3x + 4y)) \\ &\quad = (x, y), \end{aligned}$$

for each vector (x, y) in $\mathbb{R}^2$, s is the inverse function of t.

Solution to Exercise C100

(a) Since t is a linear transformation between two vector spaces of the same dimension, we use Strategy C16.

First we find a matrix representation of t. We have

$$t(1,0) = (2,4), \quad t(0,1) = (1,2).$$

Hence the matrix representation of t with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+y \\ 4x+2y \end{pmatrix}.$$

Next we evaluate the determinant of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}.$$

We have

$$\det \mathbf{A} = \begin{vmatrix} 2 & 1 \\ 4 & 2 \end{vmatrix} = 4 - 4 = 0.$$

Since $\det \mathbf{A} = 0$, t is not invertible.

(b) Since t is a linear transformation between two vector spaces of the same dimension, we use Strategy C16.

First we find a matrix representation of t. We have

$$t(1,0) = (1,3), \quad t(0,1) = (-1,1).$$

Hence the matrix representation of t with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x-y \\ 3x+y \end{pmatrix}.$$

Next we evaluate the determinant of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 3 & 1 \end{pmatrix}.$$

We have

$$\det \mathbf{A} = \begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix} = 1 - (-3) = 4.$$

Since $\det \mathbf{A} \neq 0$, t is invertible.

We now find the inverse function of t, $t^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. According to Strategy C16, t^{-1} has the matrix representation $\mathbf{v} \longmapsto \mathbf{A}^{-1}\mathbf{v}$, with respect to the standard bases for the domain and codomain. Since

$$\mathbf{A}^{-1} = \tfrac{1}{4}\begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix},$$

it follows that t^{-1} has the matrix representation

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{4}x + \frac{1}{4}y \\ -\frac{3}{4}x + \frac{1}{4}y \end{pmatrix}.$$

So t^{-1} is the linear transformation

$$\begin{aligned} t^{-1} : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longmapsto \left(\tfrac{1}{4}x + \tfrac{1}{4}y, -\tfrac{3}{4}x + \tfrac{1}{4}y\right). \end{aligned}$$

(c) Since t is a linear transformation between two vector spaces of the same dimension, we use Strategy C16.

First we find a matrix representation of t. We have

$$t(1,0,0) = (2,-1,0), \quad t(0,1,0) = (0,3,0),$$
$$t(0,0,1) = (0,0,1).$$

Hence the matrix representation of t with respect to the standard bases for the domain and codomain is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x \\ 3y - x \\ z \end{pmatrix}.$$

Next we evaluate the determinant of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We have

$$\begin{aligned} \det \mathbf{A} = \begin{vmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{vmatrix} &= 2\begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} - 0 + 0 \\ &= 2 \times 3 = 6. \end{aligned}$$

Since $\det \mathbf{A} \neq 0$, t is invertible.

We now find the inverse function of t, $t^{-1} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$. According to Strategy C16, t^{-1} has the matrix representation $\mathbf{v} \longmapsto \mathbf{A}^{-1}\mathbf{v}$, with respect to the standard bases for the domain and codomain.

Using row-reduction from Unit C1, we find

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so t^{-1} has the matrix representation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x \\ \frac{1}{6}x + \frac{1}{3}y \\ z \end{pmatrix}.$$

So t^{-1} is the linear transformation

$$\begin{aligned} t^{-1} : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto \left(\tfrac{1}{2}x, \tfrac{1}{6}x + \tfrac{1}{3}y, z\right). \end{aligned}$$

(d) Since t is a linear transformation between two vector spaces of different dimensions, it follows from Corollary C46 that t is not invertible.

Solution to Exercise C101

The linear transformation

$$\begin{aligned} t : P_3 &\longrightarrow \mathbb{R}^3 \\ a + bx + cx^2 &\longmapsto (a, b, c) \end{aligned}$$

is one-to-one and onto and hence invertible. It is therefore an isomorphism.

(There are many other possibilities.)

Solution to Exercise C102

The vector spaces $\mathbb{R}^2$, $\mathbb{C}$ and P_2 are isomorphic, since they are all two-dimensional.

The vector spaces $\mathbb{R}^3$ and P_3 are isomorphic, since they are both three-dimensional.

Solution to Exercise C103

(a) The image set of this linear transformation is the x-axis. This is a subspace of the codomain.

(b) The image set of this linear transformation is the line $y = x$. This is a subspace of the codomain.

Solution to Exercise C104

We have

$$t(1, 0) = (1, 1), \quad t(0, 1) = (0, 0).$$

The image set of t is the line $y = x$; that is,

$$\operatorname{Im} t = \{(k, k) : k \in \mathbb{R}\}.$$

Thus $\operatorname{Im} t$ is spanned by $(1, 1) = t(1, 0)$.

Solution to Exercise C105

We use Strategy C17.

(a) We take the standard basis $\{(1, 0), (0, 1)\}$ for the domain $\mathbb{R}^2$.

We determine the images of these basis vectors:

$$t(1, 0) = (1, 2), \quad t(0, 1) = (0, 1).$$

The set $\{(1, 2), (0, 1)\}$ is linearly independent, so it is a basis for $\operatorname{Im} t$.

Since the basis has two elements,

$$\dim(\operatorname{Im} t) = 2.$$

(b) We take the standard basis $\{1, x, x^2\}$ for the domain P_3.

We determine the images of these basis vectors:

$$t(1) = 0, \quad t(x) = 1, \quad t(x^2) = 2x.$$

The set $\{0, 1, 2x\}$ is not linearly independent since it contains the zero vector. We discard 0 to give the set $\{1, 2x\}$.

The vectors 1 and $2x$ are linearly independent, so $\{1, 2x\}$ is a basis for $\operatorname{Im} t$.

Since the basis has two elements,

$$\dim(\operatorname{Im} t) = 2.$$

(c) We take the standard basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for the domain $\mathbb{R}^3$.

We determine the images of these basis vectors:

$$\begin{aligned} &t(1, 0, 0) = (1, 1, 1), \quad t(0, 1, 0) = (2, 0, 1), \\ &t(0, 0, 1) = (3, 1, 2). \end{aligned}$$

The set $\{(1, 1, 1), (2, 0, 1), (3, 1, 2)\}$ is not linearly independent. In fact,

$$(3, 1, 2) = (1, 1, 1) + (2, 0, 1),$$

so we discard $(3, 1, 2)$ to give the set $\{(1, 1, 1), (2, 0, 1)\}$.

The vectors $(1, 1, 1)$ and $(2, 0, 1)$ are linearly independent, so $\{(1, 1, 1), (2, 0, 1)\}$ is a basis for $\operatorname{Im} t$.

Since the basis has two elements,

$$\dim(\operatorname{Im} t) = 2.$$

(You may have chosen to discard $(1,1,1)$ or $(2,0,1)$ instead. This would still give a correct answer.)

Solution to Exercise C106

(a) We know from Exercise C105(a) that $\dim(\operatorname{Im} t) = 2$. Thus $\operatorname{Im} t$ is the whole of the two-dimensional codomain $\mathbb{R}^2$; so t is onto.

(b) We know from Exercise C105(b) that $\dim(\operatorname{Im} t) = 2$. Thus $\operatorname{Im} t$ is the whole of the two-dimensional codomain P_2; so t is onto.

(c) We know from Exercise C105(c) that $\dim(\operatorname{Im} t) = 2$. Thus $\operatorname{Im} t$ is not the whole of the three-dimensional codomain $\mathbb{R}^3$; so t is not onto.

Solution to Exercise C107

(a) For this linear transformation, $t(x,y,z) = \mathbf{0}$ if and only if $(x,0) = (0,0)$, that is, if and only if $x = 0$. Thus the kernel of t is the (y,z)-plane. This is a subspace of the domain $\mathbb{R}^3$.

(b) For this linear transformation, $t(x,y) = \mathbf{0}$ if and only if $(x,x) = (0,0)$, that is, if and only if $x = 0$. Thus the kernel of t is the y-axis. This is a subspace of the domain $\mathbb{R}^2$.

Solution to Exercise C108

(a) The kernel of t is the set of vectors (x,y) in $\mathbb{R}^2$ that satisfy

$$t(x,y) = \mathbf{0},$$

that is,

$$(x, 2x+y) = (0,0).$$

Equating coordinates, we obtain the system

$$\begin{aligned} x \phantom{{}+y} &= 0\\ 2x + y &= 0. \end{aligned}$$

Substituting $x = 0$ from the first equation into the second equation, we obtain $y = 0$.

So the kernel of t is

$$\operatorname{Ker} t = \{(0,0)\}.$$

Since this contains only the zero vector,

$$\dim(\operatorname{Ker} t) = 0,$$

that is, $\operatorname{Ker} t$ is a zero-dimensional subspace of the domain $\mathbb{R}^2$.

(b) The kernel of t is the set of vectors (x,y,z) in $\mathbb{R}^3$ that satisfy

$$t(x,y,z) = \mathbf{0},$$

that is,

$$(x+2y+3z, x+z, x+y+2z) = (0,0,0).$$

Equating coordinates, we obtain the system

$$\begin{aligned} x + 2y + 3z &= 0\\ x \phantom{{}+2y} + z &= 0\\ x + y + 2z &= 0. \end{aligned}$$

To solve this system we row-reduce the augmented matrix.

$$\begin{array}{l} \mathbf{r}_1\\ \mathbf{r}_2\\ \mathbf{r}_3 \end{array} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0\\ 1 & 0 & 1 & 0\\ 1 & 1 & 2 & 0 \end{array}\right) \quad \begin{array}{r} 6\\ 2\\ 4 \end{array}$$

$$\begin{array}{l} \\ \mathbf{r}_2 \to \mathbf{r}_2 - \mathbf{r}_1\\ \mathbf{r}_3 \to \mathbf{r}_3 - \mathbf{r}_1 \end{array} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0\\ 0 & -2 & -2 & 0\\ 0 & -1 & -1 & 0 \end{array}\right) \quad \begin{array}{r} 6\\ -4\\ -2 \end{array}$$

$$\begin{array}{l} \\ \mathbf{r}_2 \to -\frac{1}{2}\mathbf{r}_2\\ \\ \end{array} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0\\ 0 & 1 & 1 & 0\\ 0 & -1 & -1 & 0 \end{array}\right) \quad \begin{array}{r} 6\\ 2\\ -2 \end{array}$$

$$\begin{array}{l} \mathbf{r}_1 \to \mathbf{r}_1 - 2\mathbf{r}_2\\ \\ \mathbf{r}_3 \to \mathbf{r}_3 + \mathbf{r}_2 \end{array} \quad \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0 \end{array}\right) \quad \begin{array}{r} 2\\ 2\\ 0 \end{array}$$

The augmented matrix is in row-reduced form and we have

$$\begin{aligned} x \phantom{{}+y} + z &= 0\\ y + z &= 0. \end{aligned}$$

Assigning the parameter k to the unknown z, we obtain

$$x = -k, \quad y = -k, \quad z = k.$$

So the kernel of t is

$$\operatorname{Ker} t = \{(-k,-k,k) : k \in \mathbb{R}\},$$

that is, $\operatorname{Ker} t$ is the line through $(0,0,0)$ and $(-1,-1,1)$.

Thus

$$\dim(\operatorname{Ker} t) = 1,$$

that is, $\operatorname{Ker} t$ is a one-dimensional subspace of the domain $\mathbb{R}^3$.

Solution to Exercise C109

Let $p(x) = a + bx + cx^2$ be a polynomial in P_3. Then

$$t(p(x)) = b + 2cx.$$

The kernel of t is the set of polynomials $p(x) = a + bx + cx^2$ in P_3 that satisfy

$$t(p(x)) = \mathbf{0},$$

that is,

$$b + 2cx = 0.$$

Equating coefficients, we obtain the system

$$\begin{aligned} b \quad &= 0 \\ 2c &= 0. \end{aligned}$$

So a can take any real value, $b = 0$ and $c = 0$.

Thus the kernel of t is

$$\operatorname{Ker} t = \{p(x) : p(x) = a,\ a \in \mathbb{R}\},$$

that is, the set of constant polynomials.

A basis for this subspace (the kernel) is $\{1\}$, so it follows that

$$\dim(\operatorname{Ker} t) = 1.$$

Solution to Exercise C110

(a) The kernel of t is $\operatorname{Ker} t = \{\mathbf{0}\}$. Thus t is one-to-one.

(b) The kernel of t is $\operatorname{Ker} t \neq \{\mathbf{0}\}$. Thus t is not one-to-one.

(c) The kernel of t is $\operatorname{Ker} t \neq \{\mathbf{0}\}$. Thus t is not one-to-one.

Solution to Exercise C111

(a) For the linear transformation

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x, 2x + y), \end{aligned}$$

we found in Exercise C105(a) that $\dim(\operatorname{Im} t) = 2$, and in Exercise C108(a) that $\dim(\operatorname{Ker} t) = 0$. Thus

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = 2 + 0 = 2,$$

which is the dimension of the domain $\mathbb{R}^2$.

(b) For the linear transformation

$$\begin{aligned} t : P_3 &\longrightarrow P_2 \\ p(x) &\longmapsto p'(x), \end{aligned}$$

we found in Exercise C105(b) that $\dim(\operatorname{Im} t) = 2$, and in Exercise C109 that $\dim(\operatorname{Ker} t) = 1$. Thus

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = 2 + 1 = 3,$$

which is the dimension of the domain P_3.

(c) For the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x + 2y + 3z, x + z, x + y + 2z), \end{aligned}$$

we found in Exercise C105(c) that $\dim(\operatorname{Im} t) = 2$, and in Exercise C108(b) that $\dim(\operatorname{Ker} t) = 1$. Thus

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = 2 + 1 = 3,$$

which is the dimension of the domain $\mathbb{R}^3$.

Solution to Exercise C112

(a) In this case, the dimension of the codomain (which is 3) is greater than the dimension of the domain (which is 2), so t is not onto.

(b) In this case, the codomain and the domain both have dimension 2. There are two possibilities: *either* t is both one-to-one and onto, *or* t is neither one-to-one nor onto.

(c) In this case, the dimension of the codomain (which is 2) is less than the dimension of the domain (which is 3), so t is not one-to-one.

Solution to Exercise C113

The number of solutions of this system of equations is the same as the number of vectors that map to $(1, 1, 1)$ under the linear transformation

$$\begin{aligned} t : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto (x + 2y + 3z, x + z, x + y + 2z). \end{aligned}$$

We know from the solution to Exercise C105(c) that $(1, 1, 1)$ is in the image set of t, and from

Exercise C108(b) that $\operatorname{Ker} t \neq \{\mathbf{0}\}$. Thus the system of equations has infinitely many solutions.

Solution to Exercise C114

(a) This is a system of two linear equations in three unknowns. Since $3 > 2$, the system has either no solutions or infinitely many solutions.

(b) This is a system of three linear equations in three unknowns. There are two possibilities:

- the system has exactly one solution for each set of values of a, b and c
- there are some values of a, b and c for which the system has no solutions; for all other values of a, b and c, the system has infinitely many solutions.